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

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 17, 2012

Abstract

We extend the results in the accompanying paper to include a multivariate Student's t distribution for the macroeconomic and finance variables, instead of the multivariate Gaussian distribution, that is reported in the paper. We also provide additional empirical results.

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

JEL classification codes: C32, G32.

1 A macro model for y_t^m with multivariate Student's t

Here we extend the results in the paper to include a multivariate Student's t distribution for the macroeconomic variables instead of the multivariate Gaussian distribution reported in section 3.3 of the main paper. All notation is the same as in the main paper. The log-density of the (standardized) Student's t distribution is defined as

$$\begin{aligned}\mathcal{L}_t &= \log\left[\Gamma\left(\frac{\nu+k_t}{2}\right)\right] - \log\left[\Gamma\left(\frac{\nu}{2}\right)\right] - \frac{1}{2}\log|\tilde{S}_t\Sigma_m\tilde{S}_t'| \\ &\quad -\frac{k_t}{2}\log\left[(\nu-2)\pi\right] - \frac{(\nu+k_t)}{2}\log\left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right].\end{aligned}$$

where the model is parameterized such that Σ_m is the covariance matrix. The mean of this distribution will change over time as a function of the factors.

$$\mu_t = z^m + Z^m f_t,$$

where z^m is a vector of intercepts and Z^m is a matrix of factor loadings. The GAS model requires the score vector and a scaling matrix S_t. For the scaling matrix, we will use the inverse of Fisher's information matrix. The variable k_t is the dimensionality of the observed data y_t^m.

The score with respect to f_t is

$$\nabla_t = \frac{(\nu+k_t)}{(\nu-2)}\frac{1}{w_t}\left(\tilde{S}_t Z^m\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\tilde{S}_t\left(y_t^m-\mu_t\right)$$

where we define the scalar weighting function

$$w_t = 1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}.$$

The scalar weight w_t is the key difference between the GAS model for the Student's t distribution and the Gaussian distribution. As $\nu\to\infty$, the weight converges to one and the model collapses to the Gaussian model. The same intuition applies to models with time-varying variance as shown in Creal, Koopman, and Lucas (2012) and Creal, Koopman, and Lucas (2011). The information matrix for f_t is

$$\begin{aligned}&\mathrm{E}\left[\nabla_t\nabla_t'\right] \\ &= \frac{(\nu+k_t)}{(\nu-2)}\mathrm{E}\left[\frac{1}{w_t^2}Z^{m\prime}\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(y_t-Z^m f_t\right)\left(y_t-Z^m f_t\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}Z^m\right] \\ &= \frac{(\nu+k_t)}{(\nu-2)}Z^{m\prime}\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left\{\mathrm{E}\left[\frac{1}{w_t^2}y_t y_t'\right]+\mathrm{E}\left[\frac{1}{w_t^2}\right]Z^m f_t f_t' Z^{m\prime} - 2\mathrm{E}\left[\frac{1}{w_t^2}y_t\right]f_t' Z^{m\prime}\right\}\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}Z^m.\end{aligned}$$

To compute this we define $\tilde{\nu} = \nu{+}4$ and $\tilde{\Sigma}_m = \Sigma_m \frac{\nu-2}{\nu+2}$. For the three individual expectations, we find.

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{w_t^2}\right] &= \int_{-\infty}^{\infty} \frac{1}{w_t^2} \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[(\nu-2)\pi\right]^{k_t/2}\left|\tilde{S}_t\Sigma_m\tilde{S}_t'\right|^{1/2}} \\
&\quad \left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right]^{-\frac{\nu+k_t}{2}} dy_t \\
&= \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[(\nu-2)\pi\right]^{k_t/2}\left|\tilde{S}_t\Sigma_m\tilde{S}_t'\right|^{1/2}} \int_{-\infty}^{\infty} \frac{1}{w_t^2} \\
&\quad \left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right]^{-\frac{\nu+k_t}{2}} dy_t \\
&= \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\Gamma\left(\frac{\nu+4}{2}\right)}{\Gamma\left(\frac{\nu+4+k_t}{2}\right)} \\
&= \frac{(\nu+2)\nu}{(\nu+k_t)(\nu+k_t+2)}
\end{aligned}$$

The second expectation is

$$\begin{aligned}
\mathrm{E}\left[\frac{1}{w_t^2}y_t\right] &= \int_{-\infty}^{\infty} \frac{1}{w_t^2} y_t \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[(\nu-2)\pi\right]^{k_t/2}\left|\tilde{S}_t\Sigma_m\tilde{S}_t'\right|^{1/2}} \\
&\quad \left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right]^{-\frac{\nu+k_t}{2}} dy_t \\
&= \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[(\nu-2)\pi\right]^{k_t/2}\left|\tilde{S}_t\Sigma_m\tilde{S}_t'\right|^{1/2}} \int_{-\infty}^{\infty} \frac{1}{w_t^2} y_t \\
&\quad \left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t\Sigma_m\tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right]^{-\frac{\nu+k_t}{2}} dy_t \\
&= \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\Gamma\left(\frac{\nu+4}{2}\right)}{\Gamma\left(\frac{\nu+4+k_t}{2}\right)} Z f_t \\
&= \frac{(\nu+2)\nu}{(\nu+k_t)(\nu+k_t+2)} Z^m f_t
\end{aligned}$$

The last expectation is

$$
\begin{aligned}
\mathrm{E}\left[\frac{1}{w_t^2}y_t y_t'\right] &= \int_{-\infty}^{\infty} \frac{1}{w_t^2} y_t y_t' \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\left[(\nu-2)\pi\right]^{k_t/2}\left|\tilde{S}_t \Sigma_m \tilde{S}_t'\right|^{1/2}} \\
&\quad \left[1+\frac{\left(\tilde{S}_t\left(y_t'-Z^m f_t\right)\right)'\left(\tilde{S}_t \Sigma_m \tilde{S}_t'\right)^{-1}\left(\tilde{S}_t\left(y_t-Z^m f_t\right)\right)}{(\nu-2)}\right]^{-\frac{\nu+k_t}{2}} dy_t \\
&= \frac{\Gamma\left(\frac{\nu+k_t}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{\Gamma\left(\frac{\nu+4}{2}\right)}{\Gamma\left(\frac{\nu+4+k_t}{2}\right)} \left[\tilde{\Sigma}_m + Z f_t f_t' Z^m\right] \\
&= \frac{(\nu+2)\nu}{(\nu+k_t)(\nu+k_t+2)} \left[\frac{\nu-2}{\nu+2}\Sigma_m + Z f_t f_t' Z^m\right]
\end{aligned}
$$

Combining these individual expectations into the equation above, we find the information matrix to be

$$
\mathrm{E}\left[\nabla_t \nabla_t'\right] = \frac{\nu}{(\nu+k_t+2)} \left(\tilde{S}_t Z^m\right)' \left(\tilde{S}_t \Sigma_m \tilde{S}_t'\right)^{-1} \tilde{S}_t Z^m
$$

2 Empirical results based on the (3,2,0) model

2.1 In-sample results

Here we report the empirical results based on the model (3,2,0). All factors are subject to the dynamic specification in the main paper with $p = q = 1$. Table 1 presents the estimated parameters and standard errors for this model. Standard errors are computed using the inverse Hessian of the maximized log likelihood.

The coefficients for the dynamics of the factor are strongly significant as we observed from the significance of the A coefficients in Table 1. Moreover, the macro and frailty factors are highly persistent: all of the B coefficients are estimated as 0.9 or higher. This implies that rating transition probabilities, including default probabilities, are very sticky and may deviate from their unconditional values for a substantial number of months. From a risk management perspective, this 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 estimation results for Z^m reveal that the first macro factor loads on industrial production growth, real GDP growth, the (negative) change in the unemployment rate, and to a lesser extent on the (negative) credit spread. The first three variables enter with roughly equal weights. The first macro factor can therefore be interpreted as a standard business cycle indicator. The factor is higher if the business cycle is in a good state. The results for Z^c show that the first macro factor mainly feeds significantly into the rating transition probabilities for the lower grades: bad business cycle conditions imply higher default probabilities. This holds most strongly for the CCC class, followed by the B and BB classes. The investment grade (IG) class does not load significantly on the first macro factor.

The second macro factor mainly loads on the credit spread and the equity market volatility, and to a lesser extent on the change in the unemployment rate and the negative annual stock return. All coefficients have consistent signs: high credit spreads, high volatilities, bad returns and upward changes in unemployment all push the second macro factor up. As credit spreads and volatilities are the main ingredients of the second macro factor, we interpret it as a summary of the perceptions of financial markets on economic conditions. The corresponding estimates of the elements of Z^c show that the second macro factor significantly feeds into all transition probabilities, but particularly into the transition probabilities for higher grade firms. The signs are all intuitive: if markets perceive credit risk to be high and the environment to be uncertain, we tend to see more defaults and downgrades. The second macro factor also significantly influences the mean LGD rate: high credit spreads and high volatility are positively correlated with high LGD rates.

Table 1: *Parameter estimates and standard errors for the (3,2,0) model.*

This table contains the estimated parameters and their standard errors for our model with (3,2,0) factor structure. The macros are orderd from $i = 1, \ldots, 6$ as industrial production growth (IP), unemployment rate change (UR), annual real GDP growth (GDP), credit spread (CrSPR), 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	frailty_1	frailty_2
A	0.240***	0.303***	0.278***	0.033***	0.036***
	(0.012)	(0.017)	(0.023)	(0.005)	(0.012)
B	0.968***	0.933***	0.900***	0.978***	0.973***
	(0.013)	(0.021)	(0.025)	(0.011)	(0.017)

			Z^m				
	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^m	Σ_m
IP	1.000	0.000	0.000	0.000	0.000	0.000	0.063***
	——	——	——	——	——	——	(0.006)
UR	-0.892***	0.122***	-0.062*	0.000	0.000	0.000	0.118***
	(0.037)	(0.041)	(0.040)	——	——	——	(0.012)
RGDP	0.811***	0.072	0.336***	0.000	0.000	0.000	0.278***
	(0.066)	(0.079)	(0.074)	——	——	——	(0.034)
Cr.Spr.	-0.169**	1.000	0.000	0.000	0.000	0.000	0.059***
	(0.085)	——	——	——	——	——	(0.005)
$r_{S\&P}$	0.049	-0.268***	1.223***	0.000	0.000	0.000	0.154***
	(0.093)	(0.081)	(0.093)	——	——	——	(0.012)
$\sigma_{S\&P}$	-0.007	0.648***	1.000	0.000	0.000	0.000	0.541***
	(0.107)	(0.084)	——	——	——	——	(0.045)

			Z^c					z^c	
$R_{it} =$	macro_1	macro_2	macro_3	frailty_1	frailty_2	≤IG	≤BB	≤B	≤CCC
IG	-0.052	0.202***	-0.123**	1.475***	-1.165**	6.291***	8.279***	9.617***	10.284***
	(0.059)	(0.055)	(0.069)	(0.371)	(0.555)	(0.156)	(0.175)	(0.225)	(0.276)
BB	-0.078**	0.172***	-0.102***	1.000	0.000	-5.204***	4.621***	7.047***	7.444***
	(0.037)	(0.037)	(0.040)	——	——	(0.086)	(0.082)	(0.123)	(0.140)
B	-0.184***	0.162***	-0.142***	0.970***	-0.016	-7.641***	-5.461***	4.641***	5.967***
	(0.035)	(0.031)	(0.040)	(0.156)	(0.158)	(0.135)	(0.085)	(0.080)	(0.089)
CCC	-0.262***	0.073*	-0.018	1.936***	1.000	-9.008***	-8.091***	-5.278***	3.817***
	(0.057)	(0.050)	(0.075)	(0.465)	——	(0.399)	(0.291)	(0.194)	(0.185)

	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^r	β
Z^r	0.018	0.276***	-0.082*	1.212***	1.065***	0.194*	2.569***
	(0.049)	(0.046)	(0.062)	(0.376)	(0.301)	(0.142)	(0.094)

DUM	$t = 15$	$t = 308$
IG	3.619***	-1.105
	(0.161)	(0.988)
BB	1.359***	1.813***
	(0.337)	(0.269)
B	-3.633***	-0.937***
	(0.526)	(0.281)
CCC	-3.229***	-3.058***
	(1.063)	(0.158)

The third macro factor is somewhat harder to interpret and loads on volatility, real GDP growth, and realized returns, and to a lesser extent on the (negative) change in the unemployment rate. The last four variables indicate that the third macro variable might be a proxy for business cycle conditions. However, its relation with realized volatility implies that higher uncertainty also pushes the factor up. The corresponding Z^c coefficient estimates reveal that they all have the same sign and again are predominantly significant for the high grade firms. A high value of the factor pushes default and downgrade probabilities down. The impact on the LGD rates is negative, but only significant at the 10% level. All in all this favors the interpretation of the third macro factor as another proxy for the business cycle climate.

We now turn to the frailty factors and their loadings in Z^c and Z^r. These factors explain autonomous default dynamics beyond what is implied by shared exposure to the common macro factors. The first frailty factor is the most important one and captures the default dynamics that are left after controlling for the macro factors. The coefficients for all three rating categories are substantial and significant. If the frailty factor is high, default and downgrade probabilities go up. Interestingly, the frailty factor also feeds significantly into the LGD equation. Risk of credit ratings migrations and LGD risk appear to move together.

The second frailty factor mainly loads positively on the CCC firms and negatively on IG firms. The factor captures two historical features of the corporate bond market. First, in the mid 80s and also in recent years, the number of defaults of investment grade firms is higher than expected. This causes the significantly negative coefficient on the investment grade loading in Z^c. Second, via the loading of the CCC class, the factor also picks up the benign default climate in the years preceding the financial crisis. Due to its direct link with these historical default periods, the second frailty component also impacts the LGD dynamics significantly.

Based on all estimation results, we conclude that conditioning on macro factors alone is not sufficient to capture credit risk dynamics. Both transition probabilities and LGD dynamics are affected by more than only macro factors.

The final coefficients in Table 1 are the intercepts, measurement volatilities, and dummy coefficients. The intercepts z^c of the ordered logit specification show 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 $(1 + \exp(-6.291))^{-1} \approx 99.82\%$, while the probability of a CCC company defaulting over the next month is $(1 + \exp(3.817))^{-1} \approx 2.15\%$. As expected, most of the dummy coefficients are highly significant, which is a direct consequence of the abnormally high re-rating activities in these months due to institutional features.

Table 2 presents the estimates of the parameters for the (3,2,0) model when 6 parameters of the model are restricted to be equal to zero. These parameters have been estimated initially but turned out to be insignificant. The loglikelihood for this model is -39781.8 while the AIC and BIC are 79703.6 and 80510, respectively.

Table 2: *Parameter estimates and standard errors for the (3,2,0) model with restrictions.*

This table contains the estimated parameters and their standard errors for our model with (3,2,0) factor structure. The model has 6 parameters restricted compared to the model in the paper. The macros are ordered from $i = 1, \ldots, 6$ as industrial production growth (IP), unemployment rate change (UR), annual real GDP growth (GDP), credit spread (CrSPR), 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	frailty_1	frailty_2
A	0.242***	0.305***	0.279***	0.033***	0.037***
	(0.012)	(0.016)	(0.023)	(0.005)	(0.011)
B	0.968***	0.935***	0.901***	0.977***	0.975***
	(0.014)	(0.020)	(0.025)	(0.011)	(0.017)

			Z^m				
	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^m	Σ_m
IP	1.000	0.000	0.000	0.000	0.000	0.000	0.063***
	—	—	—	—	—	—	(0.006)
UR	-0.882***	0.130***	-0.062*	0.000	0.000	0.000	0.119***
	(0.035)	(0.039)	(0.039)	—	—	—	(0.012)
RGDP	0.769***	0.000	0.333***	0.000	0.000	0.000	0.277***
	(0.050)	—	(0.071)	—	—	—	(0.034)
Cr.Spr.	-0.134**	1.000	0.000	0.000	0.000	0.000	0.059***
	(0.054)	—	—	—	—	—	(0.005)
$r_{S\&P}$	0.000	-0.299***	1.220***	0.000	0.000	0.000	0.155***
	—	(0.076)	(0.092)	—	—	—	(0.012)
$\sigma_{S\&P}$	0.000	0.638***	1.000	0.000	0.000	0.000	0.538***
	—	(0.077)	—	—	—	—	(0.045)

			Z^c				z^c		
$R_{it} =$	macro_1	macro_2	macro_3	frailty_1	frailty_2	≤IG	≤BB	≤B	≤CCC
IG	0.000	0.224***	-0.121**	1.494***	-1.100**	6.283***	8.272***	9.610***	10.276***
	—	(0.050)	(0.069)	(0.366)	(0.513)	(0.155)	(0.174)	(0.224)	(0.275)
BB	-0.067**	0.176***	-0.099***	1.000	0.000	-5.203***	4.621***	7.048***	7.444***
	(0.033)	(0.037)	(0.040)	—	—	(0.085)	(0.081)	(0.122)	(0.139)
B	-0.170***	0.168***	-0.138***	0.977***	-0.014	-7.641***	-5.460***	4.642***	5.968***
	(0.030)	(0.031)	(0.039)	(0.156)	(0.153)	(0.134)	(0.084)	(0.080)	(0.089)
CCC	-0.270***	0.077*	0.000	1.887***	1.000	-9.008***	-8.091***	-5.278***	3.818***
	(0.051)	(0.047)	—	(0.465)	—	(0.399)	(0.290)	(0.192)	(0.183)

	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^r	β_r
Z^r	0.000	0.264***	-0.071*	1.188***	1.038***	0.188*	2.569***
	—	(0.041)	(0.059)	(0.380)	(0.275)	(0.141)	(0.094)

DUM	$t = 15$	$t = 308$
IG	3.637***	-1.098
	(0.159)	(0.990)
BB	1.366***	1.808***
	(0.337)	(0.269)
B	-3.626***	-0.941***
	(0.526)	(0.281)
CCC	-3.232***	-3.058***
	(1.061)	(0.158)

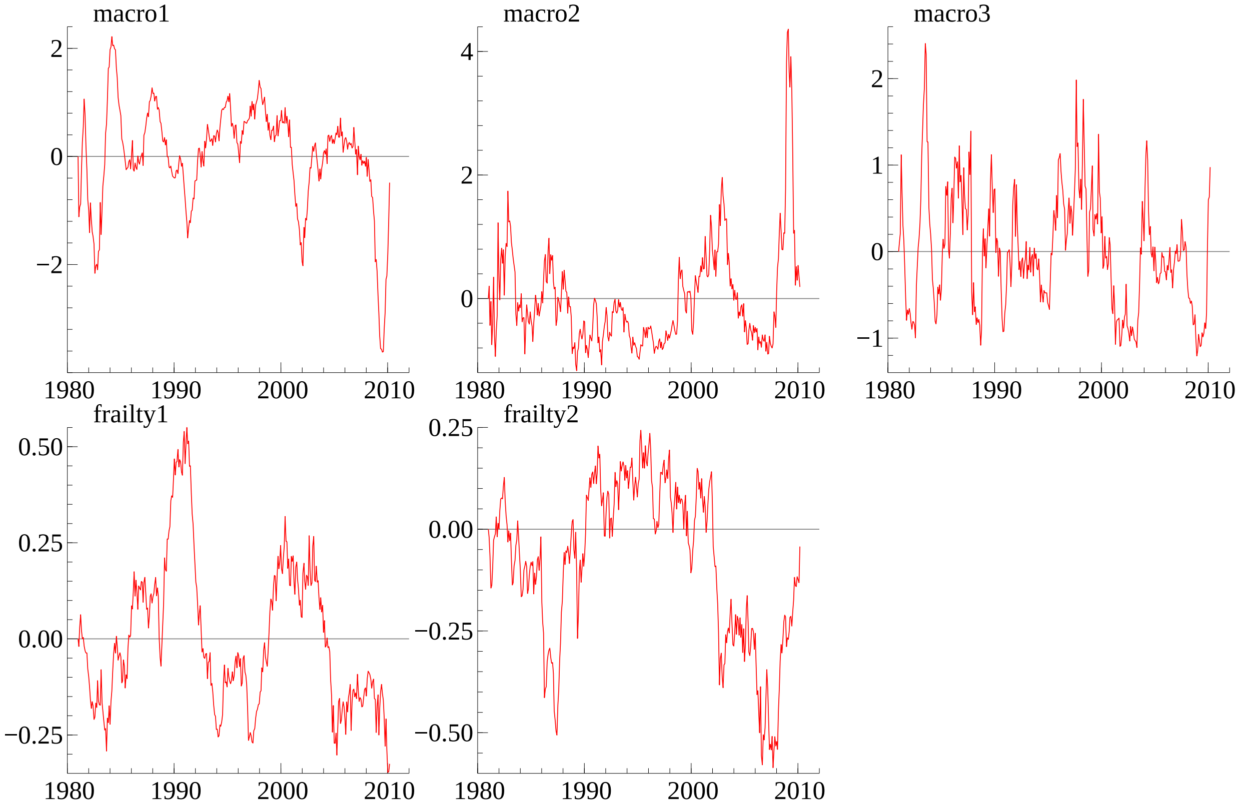

Figure 1: *Estimates of macro and frailty factors in the (3,2,0) model*

The figure contains the estimated factors for a specification with 3 macro and 2 frailty factors.

Table 3 includes estimates of the parameters of the (3,2,0) model when the macroeconomic series are modeled with a multivariate Student's t distribution with time-varying mean as above. The log-likelihood value for this model is -39777.3 and the AIC and BIC values are 79708.6 and 80596.0, respectively.

2.2 Signal extraction

Graphs of the estimated five factors are presented in Figure 1, and their implied fits for the macro series in Figure 2. We confirm our earlier finding on the relation of the first and third macro factors to growth, and of the second macro factor to credit spreads and volatilities. The fit to the macro factors is tight. Only the highest peaks in the volatility series are missed by our modeling framework. The close fit is most likely due to the dynamic factor specification of our model in combination with the monthly updating frequency of the factors.

The transition probabilities are presented in Figure 3. The probabilities are driven by both the macro and frailty factors. The first frailty factor in Figure 1 captures most of the dynamics in defaults and downgrades. In particular, we see the high peak in defaults around 1990-1991

Table 3: *Parameter estimates and standard errors for the (3,2,0) model with Student's t errors for the macro series.*

This table contains the estimated parameters and their standard errors for our model with (3,2,0) factor structure and a (standardized) multivariate Student's t distribution for the macroeconomic variables. The macros are ordered from $i = 1, \ldots, 6$ as industrial production growth (IP), unemployment rate change (UR), annual real GDP growth (GDP), credit spread (CrSPR), 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	frailty_1	frailty_2
A	0.242***	0.319***	0.279***	0.033***	0.037***
	(0.012)	(0.017)	(0.023)	(0.005)	(0.011)
B	0.970***	0.944***	0.897***	0.978***	0.973***
	(0.013)	(0.022)	(0.025)	(0.011)	(0.017)

			Z^m					
	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^m	Σ_m	ν
IP	1.000	0.000	0.000	0.000	0.000	0.000	0.062***	170.03**
	—	—	—	—	—	—	(0.006)	(71.09)
UR	-0.888***	0.125***	-0.067*	0.000	0.000	0.000	0.120***	
	(0.038)	(0.041)	(0.040)	—	—	—	(0.012)	
RGDP	0.807***	0.062	0.333***	0.000	0.000	0.000	0.279***	
	(0.066)	(0.078)	(0.075)	—	—	—	(0.034)	
Cr.Spr.	-0.172**	1.000	0.000	0.000	0.000	0.000	0.059***	
	(0.088)	—	—	—	—	—	(0.005)	
$r_{S\&P}$	0.052	-0.277***	1.222***	0.000	0.000	0.000	0.153***	
	(0.092)	(0.077)	(0.092)	—	—	—	(0.012)	
$\sigma_{S\&P}$	-0.009	0.645***	1.000	0.000	0.000	0.000	0.516***	
	(0.107)	(0.082)	—	—	—	—	(0.045)	

			Z^c					z^c	
$R_{it} =$	macro_1	macro_2	macro_3	frailty_1	frailty_2	≤IG	≤BB	≤B	≤CCC
IG	-0.053	0.192***	-0.125**	1.487***	-1.150**	6.289***	8.277***	9.615***	10.282***
	(0.059)	(0.055)	(0.070)	(0.370)	(0.535)	(0.156)	(0.175)	(0.225)	(0.276)
BB	-0.078**	0.165***	-0.102***	1.000	0.000	-5.203***	4.621***	7.047***	7.444***
	(0.036)	(0.037)	(0.040)	—	—	(0.085)	(0.082)	(0.122)	(0.139)
B	-0.184***	0.157***	-0.139***	0.977***	-0.014	-7.640***	-5.460***	4.642***	5.967***
	(0.035)	(0.031)	(0.040)	(0.157)	(0.154)	(0.134)	(0.084)	(0.080)	(0.089)
CCC	-0.259***	0.067*	-0.020	1.918***	1.000	-9.006***	-8.089***	-5.276***	3.820***
	(0.057)	(0.050)	(0.075)	(0.465)	—	(0.399)	(0.291)	(0.192)	(0.183)

	macro_1	macro_2	macro_3	frailty_1	frailty_2	z^r	β_r
Z^r	0.010	0.258***	-0.083*	1.222***	1.021***	0.190*	2.560***
	(0.048)	(0.046)	(0.062)	(0.368)	(0.286)	(0.140)	(0.094)

DUM	$t = 15$	$t = 308$
IG	3.614***	-1.105
	(0.162)	(0.989)
BB	1.349***	1.814***
	(0.338)	(0.269)
B	-3.642***	-0.932***
	(0.527)	(0.281)
CCC	-3.219***	-3.053***
	(1.065)	(0.158)

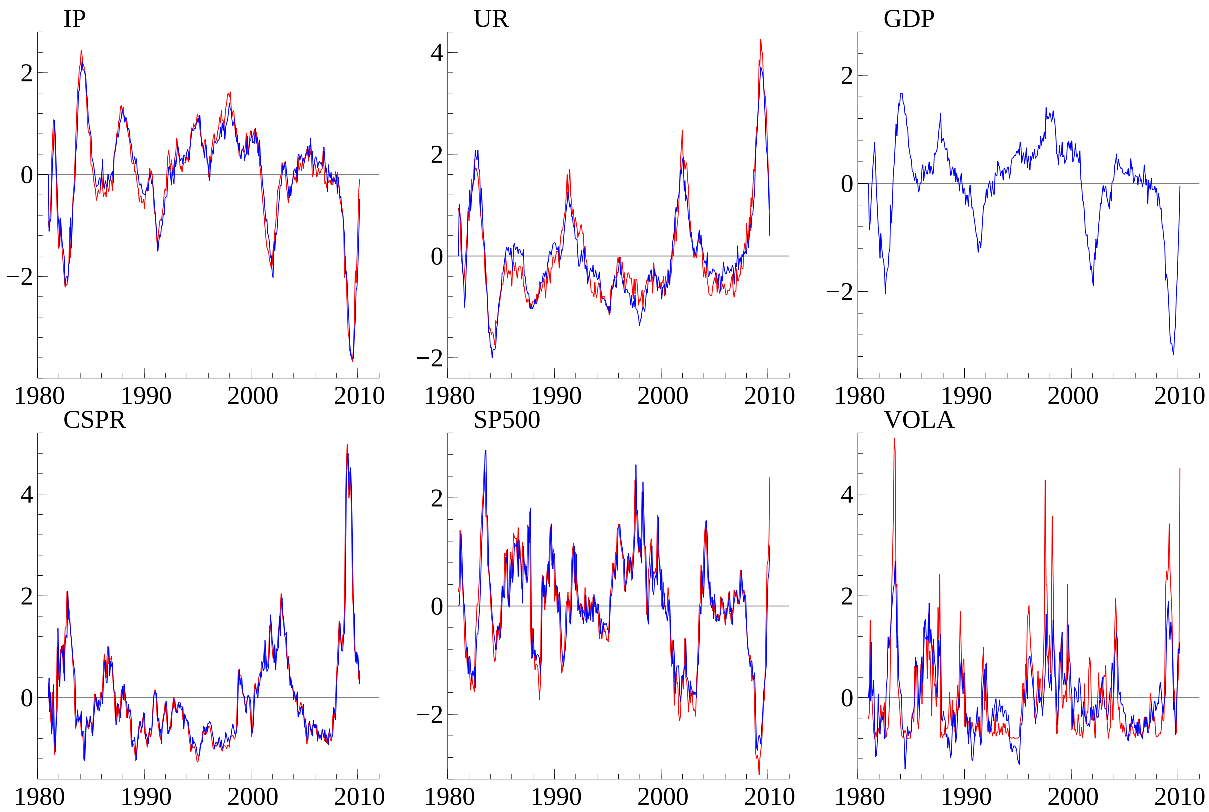

Figure 2: *fit to macro variables*

The figure presents industrial production annual growth (IP), the annual change in the unemployment rate (UR), real GDP annual growth (GDP), the credit spread (CSPR), annual return on the S&P500 index (SP500), and realized monthly volatility (VOLA) on the S&P500 (based on daily data). Each panel contains the macro series and the fit of our model with three macro and two frailty factors, (3,2,0).

and 2000-2002. Interestingly, the first frailty factor is low in the most recent years of the sample, implying that the number of defaults is lower than what should be expected based on macro fundamentals.

As mentioned before, the second frailty factor in Figure 1 captures higher downgrade and default rates during the savings and loans and the financial crisis, as well as lower default rates of CCC graded companies in the wake of the dotcom crisis. The impact of the combined macro and frailty factors on transition probabilities in Figure 3 shows the structure of the ordered logit model. In each row, the dynamics of ratings are driven by the same linear combination of unobserved factors. Given that we consider transitions at the monthly frequency, all probabilities are close to zero, except the probability that the rating remains unchanged. The peaks for the different ratings differ across rating grades. For example, the investment grade class has its highest default peaks in the financial and the dotcom crises, and some substantially lower peaks in the mid 80s and the 1991 recession. By contrast, the CCC class

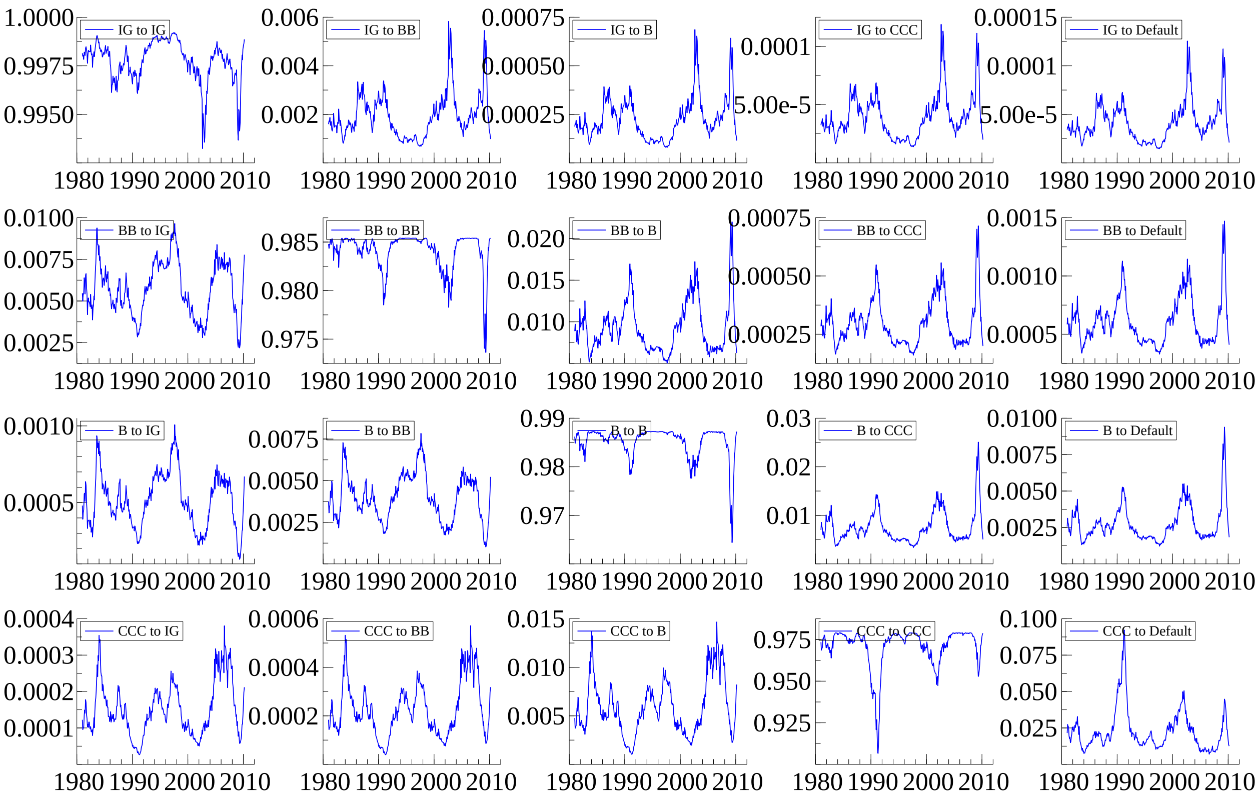

Figure 3: *Time-varying transition probabilities for the (3,2,0) model*

has its highest peak in 1991. However, even though the magnitude of the peaks and troughs differs across rating grades, all rating classes appear to share the common dynamics of clustered default experience present in the data.

Finally, we turn to the LGD series. In the upper-right panel of Figure 4 it appears that the LGD data is very noisy and irregularly spaced. During some months, we do not observe any LGDs, while in other months we observe many defaults and LGDs. The extracted factors from our model appear to represent the salient features of our data set well. LGDs correlate positively with default rate dynamics via the common factors in the model. Indeed, defaults and LGDs tend to be high and low at the same time.

The effect of the time variation in the parameters of the beta distribution is clearly visualized in the left hand panels of Figure 4. The upper-left panel gives the cross-sectional beta distributions for LGDs applicable in a mild year, June 2006. We first concentrate on the (3,2,0) model, indicated by the solid-bar histogram. In June 2006, the bulk of the probability mass for the (3,2,0) model is to the left, indicating that LGDs are typically low at that time. By contrast, the result for January 2009 in the lower-left panel are quite the opposite. Most of the probability mass has shifted to the right, indicating that LGD rates have gone up substantially

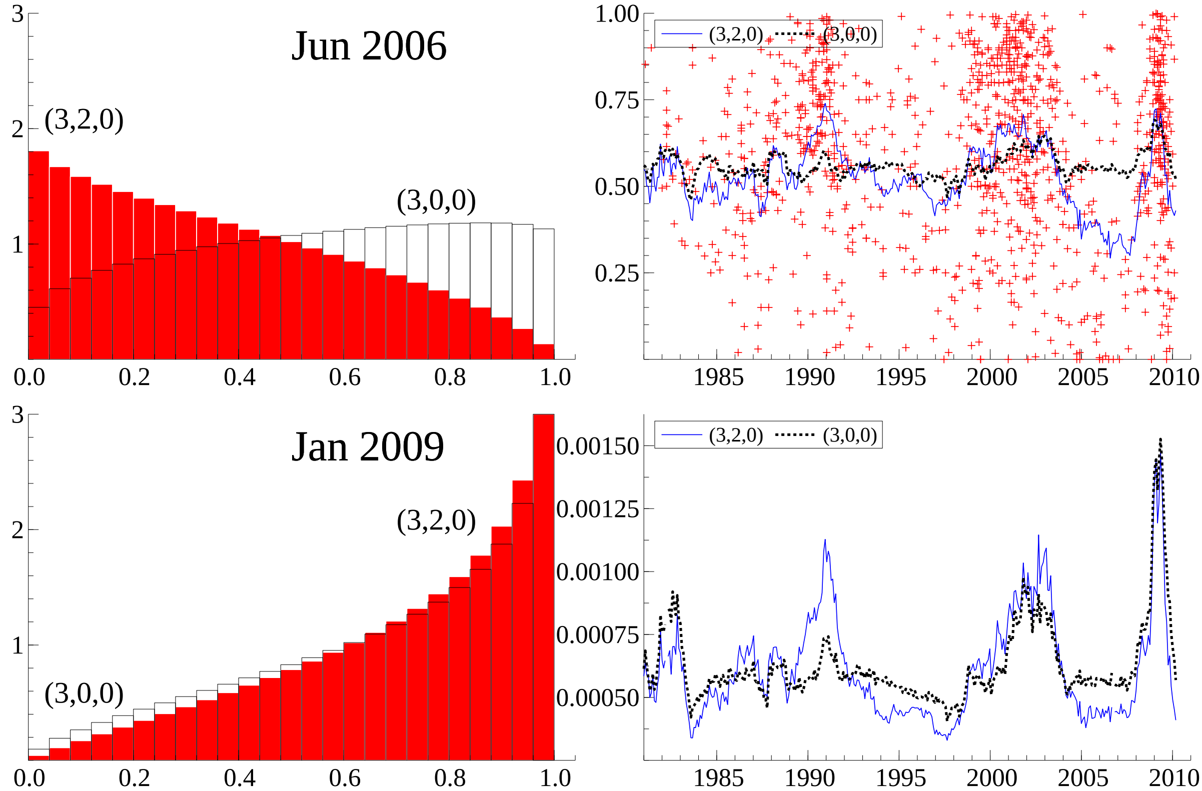

Figure 4: *Loss-Given-Default (LGD) dynamics*

The left panels contain the cross sectional beta distributions applicable in June 2006 and January 2009 for a model with three macro factors (3,0,0) and a model with three macro and two frailty factors (3,2,0). 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 (3,0,0) and (3,2,0) models.

on average.

We conclude our discussion of the in-sample results by focussing on the difference between a model with (3,2,0) and without (3,0,0) frailty effects. The upper-right graph indicates that the effect on the dynamics of LGDs is substantial. Only accounting for macro effects and discarding potential frailty, the model's fit misses both historically high and low LGD rates. This is corroborated in the left-hand graphs. During June 2006, the implied beta distribution for LGDs for the (3,0,0) is substantially different than the one for the (3,2,0) model. During other periods such as January 2009, the differences are small.

The different implications of adding frailty for credit risk are also evident if we consider transition probabilities. The lower-right panel in Figure 4 presents the transition probability from BB to default. We find that the dynamics are substantially different between the two models, and particularly that the model with only macro factors misses some of the salient peaks and troughs present in the (3,2,0) model. For example, the (3,0,0) model misses the fact that the default rates in the years leading up to the financial crises are atypically low compared to macro fundamentals.

2.3 Forecasting and credit risk management

The forecasting scenario is set up similarly as in our main paper. Figure 5 presents our initial results. It reveals the simulated loss distributions at four different horizons for the (3,2,0) model. At the one month horizon (upper-left), we clearly visualize the differences in the forecasting distribution. Starting from a recession, next month's losses are higher on average and are also more spread out. If current economic conditions are good, by contrast, next period's losses are low on average and have a smaller standard deviation. Even at the 1-month horizon, there is a significant non-overlap of the 99% confidence regions of the two densities. Both the macros and the frailty factors are in a state of recession versus expansion at time T, causing the PDs and average LGDs of the next month to be substantially different.

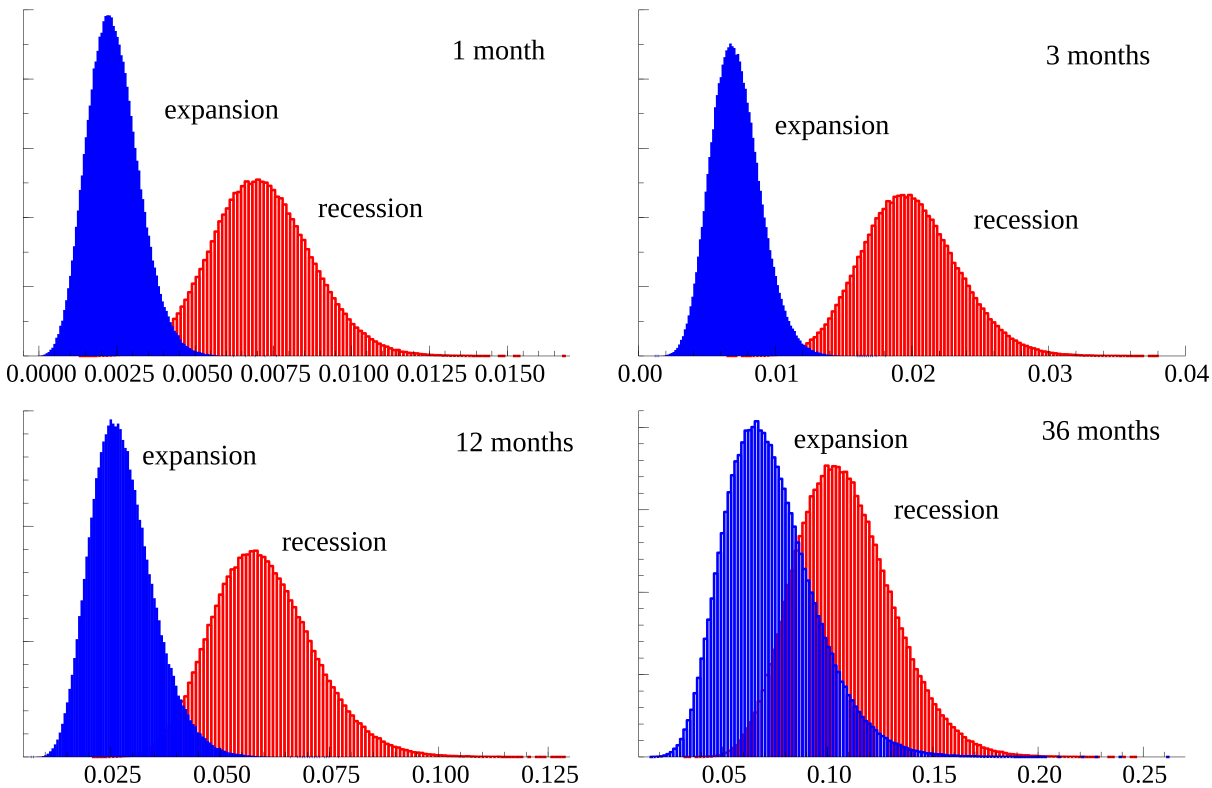

Figure 5: *Comparison of simulated loss distributions for the (3,2,0) model*

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

As the forecasting horizon H increases, the densities gradually start to overlap more and

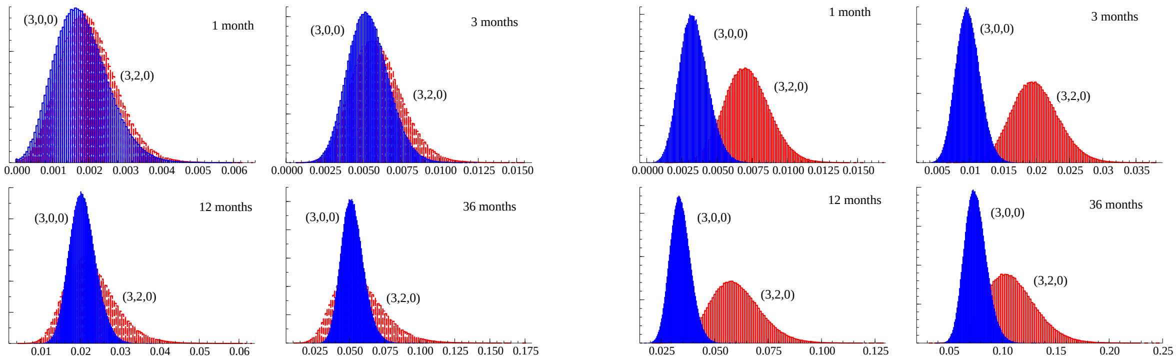

Figure 6: *Comparison of simulated loss distributions for the (3,2,0) and (3,0,0) model*

For our model with three macro and two frailty factors (3,2,0), respectively only three macro factors (3,0,0), 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.

more. The effect is due to the stationarity of the model and the high persistence of the factors in f_t. The high persistence of the unobserved components f_t causes the PDs and expected LGDs to be substantially different for a number of consecutive months. The cumulated losses over longer time periods remain substantially different. Over longer and longer horizons, however, the stationarity of the model (mean reversion) takes over and the influence of initial conditions starts to vanish. This can be viewed at a horizon of three years. Over a 36 months horizon, a current recession might easily turn into an expansionary phase, and vice versa.

We compare the economic contribution of the frailty factors to capital requirements in Figure 6. The same approach is followed as for Figure 5, but for the (3,2,0) and the (3,0,0) model, respectively. For the left-hand four panels in the figure, the factors are started at their unconditional means, $f_t = 0$. The panels show that the forecasting distributions of models with and without the two frailty factors are roughly similar at the one-month horizon. If anything, the probability of large losses is somewhat larger for the model with frailty factors (3,2,0).

As we start to focus on longer horizons, the differences become much clearer. The models that only use the three macro factors to drive the credit loss conditions result in much more concentrated loss distributions. The location of these distributions is roughly the same as that of the model with the frailty components, but the spread and right hand tail are substantially different. This is most clearly seen at the 36-month horizon: the (3,2,0) model has the right skew that is typical for portfolio loss distributions. Because of the additional, separate dynamics of the frailty components versus the macro components, the 99% confidence loss quantiles for the (3,2,0) are substantially larger than for the (3,0,0) model. A model that only conditions on macro risk leads to a severe underestimation of required capital, particularly at longer horizons.

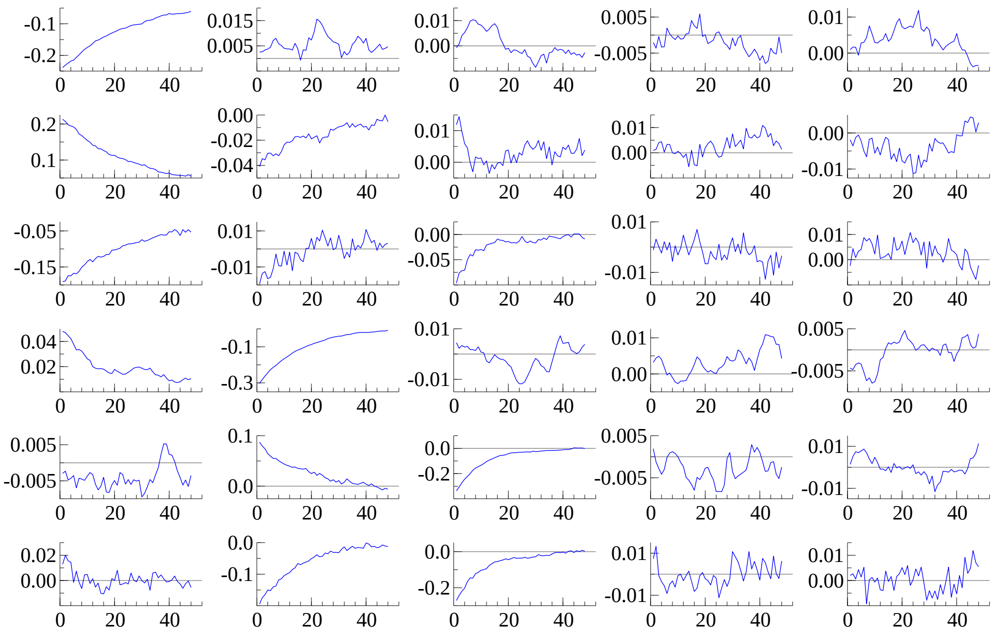

Figure 7: *Non-linear impulse response functions for the (3,2,0) model: macros*

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 5 for shocking the macro factors 1 to 3 and the 2 frailty factors. Factors are started at their mean values $f_T = 0$.

The four right-hand panels in Figure 6 give the results if the factors f_T are started in a recession. At short horizons, the results are now much more pronounced. This is due to the fact that not only the macro factors start under bad economic circumstances but the frailty factors do as well. The projected increase in expected loss rates are substantial. The upper quantiles of the loss distribution increase substantially if the frailty factors are added. The increase for the highest quantiles is close to 100% in most settings, implying a doubling of model based capital requirements at these horizons. At the longest horizon of 36 months, the distributions start to overlap more, similar to the setting with $f_T = 0$. This is again due to the stationarity of the model: at these longer horizons, the differences in initial conditions matter less. Still, the loss quantiles are substantially different for models with and without frailty.

2.4 Impulse response analysis

The impulse repsonse functions are computed as in our main paper. The impulse response functions for the macros confirm the estimation results as presented in Table 1. The impulse response function for the first macro factor (as presented in the first column of figures) has

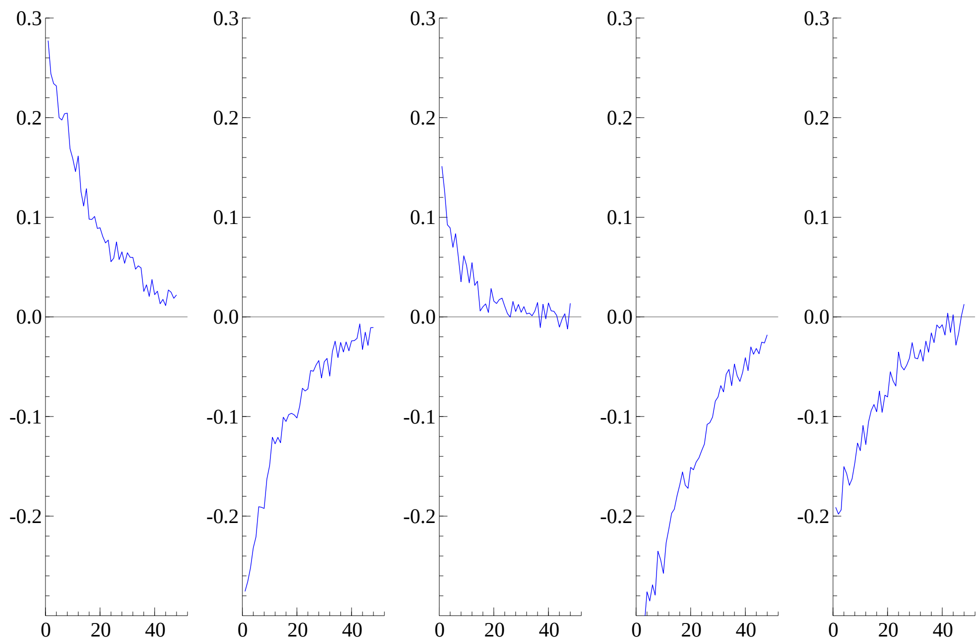

Figure 8: *Non-linear impulse response functions for the (3,2,0) model: portfolio losses*

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 simulation average of the simulated quantity for a unit size shock to one of the factors and the simulation 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 99th 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 5 are for a shock to the macro factors 1 to 3 and the 2 frailty factors, respectively. Factors are started at their values fitted at the end of our sample.

the largest impact on the business cycle related variables. A bad shock decreases industrial production growth (1st row) and real GDP growth (3rd row) and increases the unemployment rate (2nd row). The impact is quite persistent and roughly needs 3 to 4 years to die out. The second factor mainly impacts the lower three macros: the credit spread, the stock market return, and the volatility. Again, the effect of a shock in the second factor dies out only slowly: its impact lasts 3 to 4 years. The third factor appears to be of a much more transitory nature and impacts the stock market return and its volatility (and to a lesser extent real GDP growth).

As expected, the frailty factors have no direct impact on the macro variables. This underlines the recursive structure of the model and helps to interpret the frailty factors as credit dynamics above and beyond macro developments.

The results for the credit loss distributions and their 99% quantiles are in Figure 8. The figures for the 99% quantiles are more noisy due to simulation error. Both the means and quantiles, however, reveal the same pattern. We see that the first and second macro factor have the largest and most persistent impact on portfolio credit losses. The effect of a shock in the first or second macro factor can last up to three or four years. The third macro factor has a much shorter half-life and vanishes after one to two years. Its impact on portfolio credit losses is also substantially smaller.

The frailty factors have a large impact on credit losses. The first frailty factor (column 4) has an effect similar in size to the business cycle related macro factor (column 1) and the financial markets related macro factor (column 2). The impact of the second frailty factor is roughly half this size. This implies that by omitting the frailty component from the model, one may miss one third to one half of the credit risk. In addition, the dynamics of credit losses might be completely misspecified if the frailty components are left out of the model.

References

Azizpour, S., K. Giesecke, and G. Schwenkler (2010). Exploring the sources of default clustering. *Stanford University working paper series*.

Creal, D. D., S. J. Koopman, and A. Lucas (2011). A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business & Economic Statistics 29*(4), 552–563.

Creal, D. D., S. J. Koopman, and A. Lucas (2012). Generalized Autoregressive Score Models with Applications. *Journal of Applied Econometrics*.

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.

Duffie, D., A. Eckner, G. Horel, and L. Saita (2009). Frailty correlated default. *Journal of Finance 64*(5), 2089–2123.

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., 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.