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

Drew Creal[a], Bernd Schwaab[b]
Siem Jan Koopman[c,e], André Lucas[d,e]

[a]University of Chicago, Booth School of Business
[b]European Central Bank
[c]Department of Econometrics, Vrije Universiteit Amsterdam
[d]Department of Finance, Vrije Universiteit Amsterdam, and Duisenberg school of finance
[e]Tinbergen Institute, Amsterdam

Score Workshop, Amsterdam, Jan. 17-18, 2013

Motivation

- Economic time series often share common features, e.g. business cycle dynamics.
- Economic time series may be continuous and/or discrete and be observed at different frequencies.
- Credit risk application:
  1. Credit conditions depend on the macroeconomy
  2. Corporate defaults tend to "cluster" together in time.
  3. Default probabilities are higher than can be explained by covariates.

Contributions

- We introduce observation-driven mixed measurement panel data models.
- The approach allows for non-linear, non-Gaussian models with common factor across different distributions.
- Credit risk application: we develop new models for credit ratings transitions and loss-given-default (LGDs).
- The models include:
    1. Time-varying ordered logit
    2. Time-varying beta distribution

Mixed measurement panel data models

$$
\begin{aligned}
y_{it} &\sim p_i(y_{it}|F^t, Y^{t-1}; \psi), \quad i = 1, \ldots, N, \\
f_{t+1} &= \omega + B_1 f_t + A_1 s_t
\end{aligned}
$$

with log-likelihood function

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

- The observation vector y_t is $N \times 1$.
- The individual observations y_{it} may come from different distributions.
- KEY: The factors f_t may be common across distributions.
- δ_{it} is an indicator function equal to 1 if y_{it} is observed and zero otherwise. Missing values are naturally taken into account.

Mixed measurement panel data models

$$
\begin{aligned}
y_{it} &\sim p_i(y_{it}|F^t, Y^{t-1};\psi), \quad i = 1,\dots,N, \\
f_{t+1} &= \omega + B_1 f_t + A_1 s_t
\end{aligned}
$$

The score function is

$$
\begin{aligned}
s_t &= S_t \nabla_t \\
\nabla_t &= \sum_{i=1}^{N} \delta_{it} \nabla_{i,t} = \sum_{i=1}^{N} \delta_{it} \frac{\partial \log p_i(y_{it}|F^t, Y^{t-1};\psi)}{\partial f_t},
\end{aligned}
$$

- KEY: The score function allows us to pool information from different observations to estimate the common factor f_t.
- The score function also allows us to "weight" the information coming from different types of data.

Scaling matrix

Consider the eigenvalue-eigenvector decomposition of Fisher's (conditional) information matrix

$$\mathcal{I}_t \quad = \quad E_{t-1}[\nabla_t \nabla_t'] \; = \; U_t \Sigma_t U_t',$$

We define the scaling matrix as

$$S_t \quad = \quad U_t \Sigma_t^{-1/2} U_t'$$

- S_t is then the "square root" of a generalized inverse.
- The innovations s_t driving f_t have an identity covariance matrix, when the info. matrix is non-singular.
- The conditional information matrix is additive for our models:

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

Log-likehood function and ML estimation

- The log-likelihood function for an observation-driven model can easily be computed.
- The ML estimator is

$$\hat{\psi} \quad = \quad \arg\max_{\psi} \sum_{t=1}^{T} \sum_{i=1}^{N} \delta_{it} \log p_i(y_{it}|F^t, Y^{t-1}; \psi),$$

- Estimation is similar to a GARCH model.
- For a given value of ψ, the factors are computed $\{f_0, f_1, \ldots, f_T\}$ from the recursion:

$$f_{t+1} \quad = \quad \omega + B_1 f_t + A_1 s_t$$

Credit risk

- Growing econometrics literature on models for credit risk: McNeil et al. (2005), Bauwens and Hautsch (JFEct, 2006), Gagliardini and Gourieroux (JFEct, 2005), Koopman Lucas and Monteiro (JEct, 2008), Duffie et al. (JFE, JoF 2008).
- Basic observations:
  1. Probability of default varies over time with the business cycle.
  2. Conditional on default, the loss (recovery rate) varies with the business cycle.
  3. We observe excess clustering of defaults and ratings transitions beyond what can be explained by simply adding covariates.
  4. The literature focuses on a credit risk or frailty factor.
- Industry standard models are too simple to capture these features.
- New models in the literature are parameter driven models requiring simulation methods for estimation.
- We provide observation driven alternatives.

Data: Moody's and FRED

- We observe data from Jan. 1980 to March 2010.
- $7,505$ companies are rated by Moody's.
- We pool these into 5 ratings categories (IG, BB, B, C, D).
- We observe transitions, e.g. IG $\rightarrow$ BB or C $\rightarrow$ D
- There are $J = 16$ total types of transitions.
- 19,450 total credit rating transitions.
- 1,342 transitions are defaults.
- 1,125 measurements of loss-given default (LGD).
- LGD is the fraction of principal an investor loses when a firm defaults.
- We also observe six macroeconomic variables: industrial production growth, credit spread, unemployment, annual S&P500 returns, realized volatility, real GDP growth (qtly).

Models

- Credit ratings can be modeled using the (static) ordered probit model of CreditMetrics; one of the current industry standards, see Gupton Stein (2005).
- LGD's are often modeled by (static) beta distributions.
- GOAL: Build models that improve on current industry standards and are (relatively) easy to implement and estimate.

  1. Time-varying ordered logit
  2. Time-varying beta distribution

- Forecasting credit risk.
- Simulation of loss distributions and scenario analysis.
- Bank executives and regulators and can use them for "stress testing."

Mixed measurement model for credit risk

$$\begin{aligned} y_t^m &\sim \mathrm{N}\left(\mu_t, \Sigma_m\right) \\ y_{i,t}^c &\sim \text{Ordered Logit}\left(\pi_{ijt}, j \in \{\text{IG, BB, B, C, D}\}\right), \\ y_{k,t}^r &\sim \text{Beta}\left(a_{kt}, b_{kt}\right), \qquad k = 1, \ldots, K_t, \end{aligned}$$

- y_t^m are the macro variables.
- $y_{i,t}^c$ are indicator variables for each credit rating j for firm i.
- $y_{k,t}^r$ are the LGDs for the k-th default.
- K_t are the number of defaults in period t.
- μ_t, π_{ijt}, and (a_{kt}, b_{kt}) are functions of an $M \times 1$ vector of factors f_t.

Time varying Gaussian model for macro data

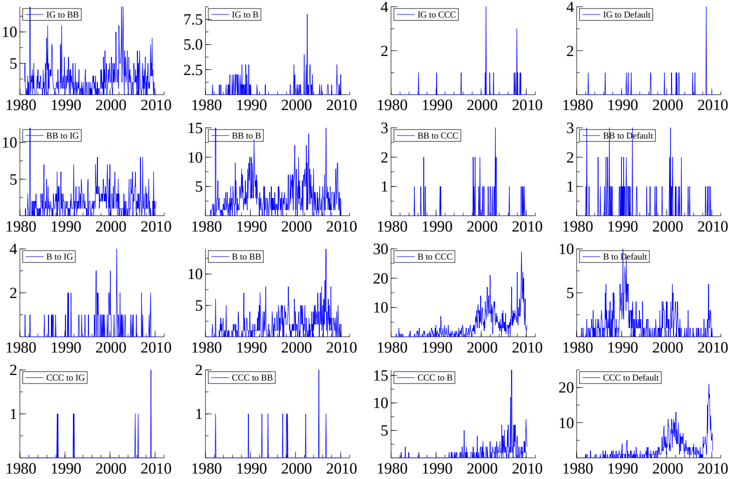
$$
\begin{aligned}
y_t^m &\sim \mathrm{N}\left(\mu_t, \Sigma_m\right), \\
\mu_t &= Z^m f_t.
\end{aligned}
$$

- Z^m is a $(6 \times M)$ matrix of factor loadings.
- Σ_m is a (6×6) diagonal covariance matrix.
- $\tilde{S}_t$ is a selection matrix indicating which macro variables are observed at time t.

$$
\begin{aligned}
\nabla_t^m &= \left(\tilde{S}_t Z^m\right)' \left(\tilde{S}_t \Sigma_m \tilde{S}_t'\right)^{-1} \tilde{S}_t \left(y_t^m - \mu_t\right), \\
\mathcal{I}_t^m &= \left(\tilde{S}_t Z^m\right)' \left(\tilde{S}_t \Sigma_m \tilde{S}_t'\right)^{-1} \tilde{S}_t Z^m.
\end{aligned}
$$

Moody's monthly credit ratings transitions

The data have been pooled together each month.



Time-varying ordered logit

$$
\begin{aligned}
y^c_{i,t} &\sim \text{Ordered Logit}\left(\pi_{ijt}, j \in \{\text{IG, BB, B, C, D}\}\right), \\
\pi_{ijt} &= \mathrm{P}\left[R_{i,t+1} = j\right] = \tilde{\pi}_{ijt} - \tilde{\pi}_{i,j-1,t}, \\
\tilde{\pi}_{ijt} &= \mathrm{P}\left[R_{i,t+1} \leq j\right] = \frac{\exp(\theta_{ijt})}{1 + \exp(\theta_{ijt})}, \\
\theta_{ijt} &= z^c_{ijt} - Z^{c\prime}_{it} f_t.
\end{aligned}
$$

- y^c_{it} is an indicator variable for each rating type.
- $J^c = 5$ categories $j \in \{\text{IG, BB, B, C, D}\}$.
- R_{it} is the rating for firm i at the start of month t.
- π_{ijt} is the probability that firm i is in category j.
- $\tilde{\pi}_{i,\text{D},t} = 0$ and $\tilde{\pi}_{i,\text{IG},t} = 1$.
- To our knowledge, a time-varying ordered logit model is new.

Time-varying ordered logit

The contribution to the log-likelihood at time t is

$$\ln p_i(y_{it}^c | F^t, Y^{t-1}; \psi) \quad = \quad \sum_{i=1}^{N_t} \sum_{j=1}^{J^c} y_{ijt}^c \log\left(\pi_{ijt}\right)$$

The score and information matrices are

$$\begin{aligned} \nabla_t^c &= -\sum_{i=1}^{N_t} \sum_{j=1}^{J^c} \frac{y_{ijt}^c}{\pi_{ijt}} \cdot \dot{\pi}_{ijt} \cdot Z_{it}^c, \\ \mathcal{I}_t^c &= \sum_{i=1}^{N_t} n_{it} \left(\sum_j \frac{\dot{\pi}_{ij,t}^2}{\pi_{ij,t}} \right) Z_{it}^c Z_{it}^{c\prime} \end{aligned}$$

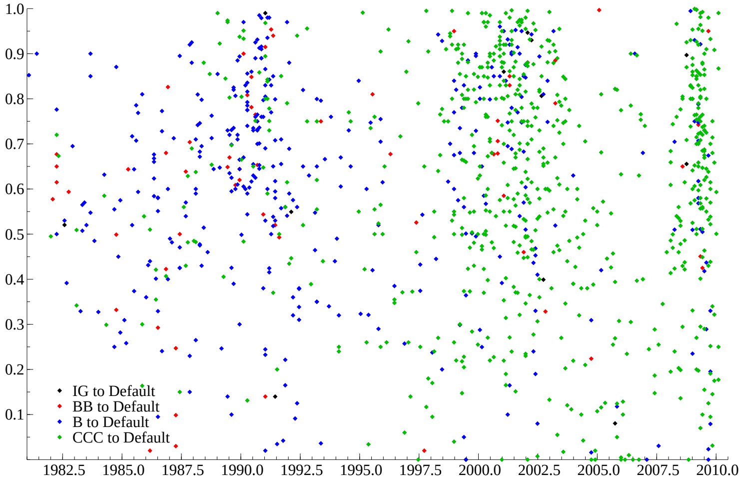
where

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

Loss given default

- When a firm defaults, investors typically lose a fraction of their investment (alternatively, they recover a fraction of their investment).
- The fraction of losses experienced by investors also varies with the business cycle.
- We develop a new model for a time-varying beta distribution.
- See McNeil and Wendin (2007 JEmpFin) for Bayesian inference in a state space model.

Loss given default by transition type



Time-varying beta distribution

$$
\begin{aligned}
y^r_{k,t} &\sim \text{Beta}\left(a_{kt}, b_{kt}\right), \qquad k = 1, \ldots, K_t, \\
a_{kt} &= \beta_r \cdot \mu^r_{kt} \\
b_{kt} &= \beta_r \cdot (1 - \mu^r_{kt}) \\
\log\left(\mu^r_{kt} / (1 - \mu^r_{kt})\right) &= z^r + Z^r f_t.
\end{aligned}
$$

- We observe $K_t \geq 0$ defaults at time t.
- $0 < y^r_{k,t} < 1$ is the amount lost conditional on the k-th default.
- μ^r_{kt} is the mean of the beta distribution.
- z^r is the unconditional level of LGDs.
- Z^r is a $(1 \times M)$ vector of factor loadings.
- β_r is a scalar parameter

Time-varying beta distribution

The contribution to the log-likelihood at time t is

$$\begin{aligned}\ln p_i(y^r_{kt}|F^t, Y^{t-1};\psi) \quad = \quad & \sum_{k=1}^{K_t}(a_{kt}-1)\log\left(y^r_{kt}\right)+(b_{kt}-1)\log\left(1-y^r_{kt}\right) \\ & -\log\left[B\left(a_{kt},b_{kt}\right)\right]\end{aligned}$$

The score and information matrices are

$$\begin{aligned}\nabla^r_t \quad &= \quad \beta_r\sum_{k=1}^{K_t}\mu^r_{kt}(1-\mu^r_{kt})\left(Z^r\right)'(1,-1)\left((\log(y^r_{kt}),\log(1-y^r_{kt}))'-\dot{B}\left(a_{kt},b_{kt}\right)\right) \\ \mathcal{I}^r_t \quad &= \quad \beta_r\sum_{k=1}^{K_t}(\mu^r_{kt}(1-\mu^r_{kt}))^2\left(Z^r\right)'(1,-1)\left(\ddot{B}\left(a_{kt},b_{kt}\right)\right)(1,-1)'\,Z^r\end{aligned}$$

where

$$\sigma^2_{kt}=\mu^r_{kt}\cdot(1-\mu^r_{kt})/(1+\beta_r).$$

Estimation details

- The macro data y_t^m has been standardized.
- We consider models with $p = 1$ and $q = 1$ factor dynamics.
- For identification of the level parameters, we set $\omega = 0$ in the factor recursion:

$$f_{t+1} = A_1 s_t + B_1 f_t$$

- For identification of the factors, we also impose restrictions on Z^m, Z^c, and Z^r.
- Some parameters have been pooled for "rare" transitions; e.g., IG $\rightarrow$ D and BB $\rightarrow$ D.
- Moody's re-defined several categories in April 1982 and Oct. 1999 causing incidental re-ratings (outliers), which we handle via dummy variables for these dates.

AIC, BIC, and log-likelihoods for different models

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

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

The number of factors for each data type are represented by (m, c, r).

Parameter estimates for the (4,2,0) model

Factor dynamics and macro loadings Z^m

	macro$_1$	macro$_2$	macro$_3$	macro$_4$	frailty$_1$	frailty$_2$
A	0.221***	0.154***	0.300***	0.282***	0.033***	0.036***
	(0.011)	(0.014)	(0.016)	(0.023)	(0.006)	(0.010)
B	0.966***	0.974***	0.924***	0.896***	0.974***	0.981***
	(0.012)	(0.016)	(0.022)	(0.026)	(0.012)	(0.012)

Z^m

	macro$_1$	macro$_2$	macro$_3$	macro$_4$	frailty$_1$	frailty$_2$
IP	1.000	0.000	0.000	0.000	0.000	0.000
	—	—	—	—	—	—
UR	-0.467***	1.000	0.000	0.000	0.000	0.000
	(0.061)	—	—	—	—	—
RGDP	0.675***	-0.266**	0.000	0.295***	0.000	0.000
	(0.070)	(0.120)	—	(0.068)	—	—
Cr.Spr.	-0.275***	0.000	1.000	0.000	0.000	0.000
	(0.075)	—	—	—	—	—
$r_{S\&P}$	0.000	-0.358**	-0.293***	1.179***	0.000	0.000
	—	(0.157)	(0.076)	(0.086)	—	—
$\sigma_{S\&P}$	0.101*	0.245*	0.563***	1.000	0.000	0.000
	(0.096)	(0.176)	(0.084)	—	—	—

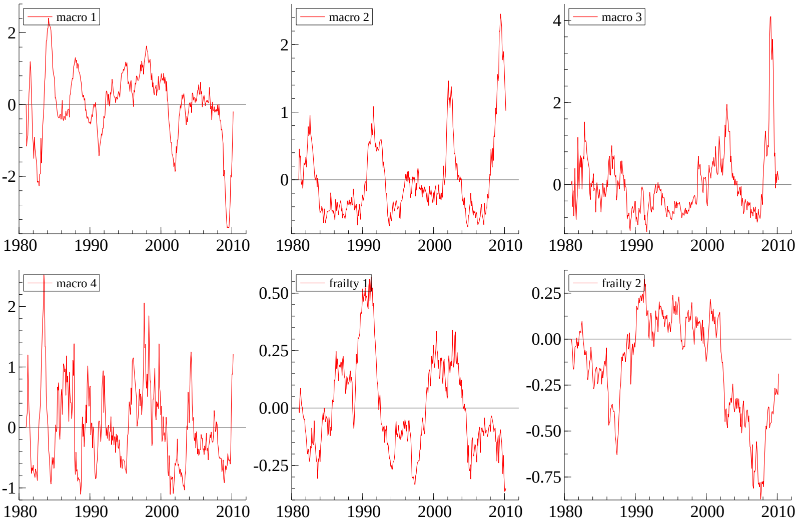
Parameter estimates for the (4,2,0) model

Credit rating and LGD loadings Z^c and Z^r

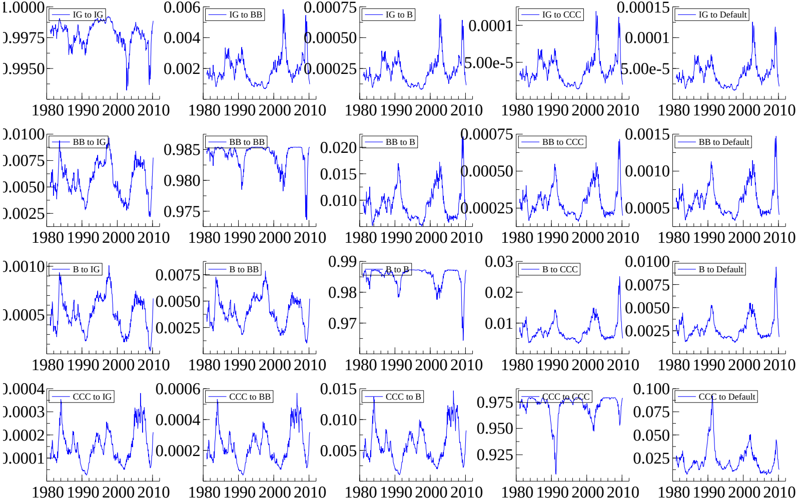
Z^c	macro_1	macro_2	macro_3	macro_4	frailty_1	frailty_2
IG	-0.064*	0.000	0.217***	-0.110*	1.520***	-0.727**
	(0.051)	—	(0.053)	(0.064)	(0.283)	(0.270)
BB	0.000	0.204***	0.158***	-0.077*	1.000	0.000
	—	(0.057)	(0.038)	(0.041)	—	—
B	-0.154***	0.130**	0.150***	-0.121***	0.914***	0.000
	(0.036)	(0.056)	(0.031)	(0.038)	(0.137)	—
CCC	-0.283***	0.000	0.076*	0.000	1.486***	1.000
	(0.052)	—	(0.048)	—	(0.418)	—

	macro_1	macro_2	macro_3	macro_4	frailty_1	frailty_2
Z^r	0.000	0.000	0.274***	-0.077*	0.938**	0.913***
	—	—	(0.045)	(0.057)	(0.315)	(0.199)

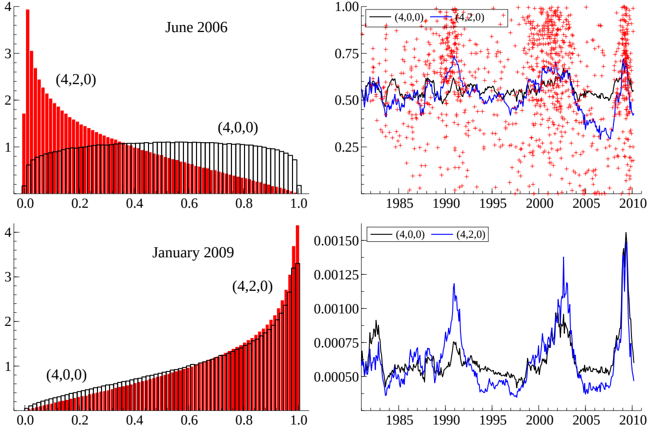
Estimated factors for the (4,2,0) model



Time-varying transition probabilities



Loss given default results



Top and bottom left are loss distributions. Top right is a plot of the mean through time. Bottom right are transition probabilities from BB $\rightarrow$ D.

Simulating cumulative loss distributions

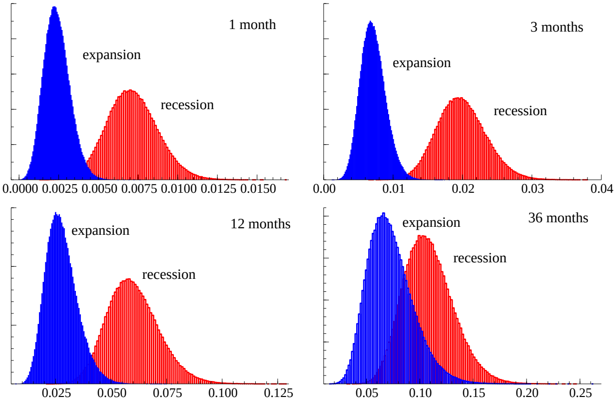
- Most financial institutions carry a large portfolio of credit related securities.
- Given a portfolio at time T, we can use the models to simulate different possible risk scenarios.
- GOAL: determine the amount of capital banks may need in the future.
- What happens if we do not include time-varying parameters f_t in the model?
- Scenario analysis:
  1. What happens if there is a negative shock to RGDP?
  2. What happens if there is an increase to credit spreads?

Simulating cumulative loss distributions

- At time T, a financial institution holds a portfolio of bonds.
- The goal is to forecast the loss distribution at time $T + h$.
- We assume a portfolio of firms with 1144 firms rated IG, 265 firms rated BB, 615 firms rated B, and 311 firms rated CCC.
- In the paper, we consider losses due only to default.
- For simplicity, we do not assume a time-varying discount function.
- We use $500,000$ simulations.
- We can start at different values f_T.

Simulating cumulative loss distributions

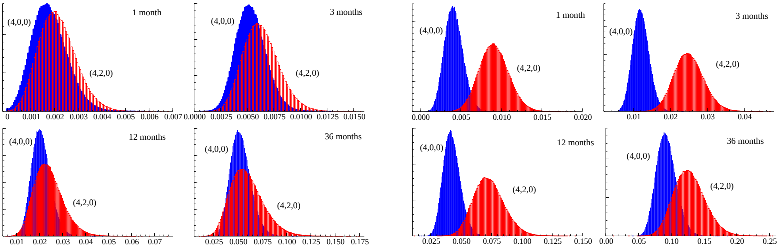
Cumulative losses on a portfolio of bonds at different horizons.



Comparison between different values of f_T starting in a recession and expansion for the (4,2,0) model.

Simulating cumulative loss distributions

Comparison of cumulative loss distributions with/without factors.



Left: starting at $f_T = 0$. Right: f_T starting in a recession.

Conclusion and future work

- We introduce a new class of observation-driven models for mixed-measurement data which share exposure to common factors.
- Missing values and mixed frequencies are handled in a natural way.
- Using this approach, we develop new models for credit risk.
- The models can be used for simulating loss distributions, stress testing, and scenario analysis.
- Future work:
  - When computing loss distributions, current models do not account for changes in market prices of bonds or loans.
  - Current models depend on industry credit ratings by Moody's, Fitch, Standard & Poors.
  - Potential to use alternative sources of data.