Recovery Rates, Default Probabilities and the Credit Cycle

Max Bruche and Carlos González-Aguado

CEMFI

BIS Stress Testing workshop, March 2007

Motivation


default rate
0.5
2.0
3.5
avg. recovery rate
30
50
1985
1990
1995
2000
2005

The research questions & method

The questions:

- Can you exploit the joint time-series behaviour of default rates and recovery rates to characterize and forecast credit risk?
- How bad is it to treat recovery rates as constant (or independent of default probabilities)?

The method:

- We propose an econometric model in which both the time variation in default probabilities and recovery rate distributions is driven by an unobserved Markov chain, the "credit cycle".

Related literature

- Default probabilities or rating transitions vary over time, and are related to macro variables.
  (Bangia et al. (2002), Nickell et al. (2000))
- Recovery rates and default probabilities are contemporaneously negatively related.
  (Altman et al. 2006, Acharya et al. 2007)
- Recovery rates and default probabilities can be modelled as functions of observed covariates.
  (Chava et al. 2006)
- Theory:
  - Recovery rates should be related to the state of the industry: Shleifer and Vishny (1992).
  - RBC and credit: Bernanke and Gertler (1989), Williamson (1987)

Summary of results

Summary of results

Observed covariates versus latent factor:

- The proposed model describes the data well, and does better than many models based on observed covariates.
- E.g. out-of-sample rolling RMSE for predicted recovery rates is 22.86%. (Chava et al. 2006: 24.96%)

Summary of results

Observed covariates versus latent factor:

- The proposed model describes the data well, and does better than many models based on observed covariates.
- E.g. out-of-sample rolling RMSE for predicted recovery rates is 22.86%. (Chava et al. 2006: 24.96%)

We can use the estimated model to look at what happens when we go from constant to time-varying recovery rate distributions. We get

- higher estimates of tail risk,
  (for a sample portfolio, the 99% VaR goes from 2.6% to 2.9%.)
- the same expected losses,
- bigger swings in spreads over the cycle. Average spreads over the cycle are not affected.

Summary of results

Observed covariates versus latent factor:

- The proposed model describes the data well, and does better than many models based on observed covariates.
- E.g. out-of-sample rolling RMSE for predicted recovery rates is 22.86%. (Chava et al. 2006: 24.96%)

We can use the estimated model to look at what happens when we go from constant to time-varying recovery rate distributions. We get

- higher estimates of tail risk,
  (for a sample portfolio, the 99% VaR goes from 2.6% to 2.9%.)
- the same expected losses,
- bigger swings in spreads over the cycle. Average spreads over the cycle are not affected.

Caveats?

The basic idea

The DGP works as follows:

- The state of the credit cycle is determined by a two-state Markov chain.
- The number of defaulting firms is drawn using the state-dependent default probability.
- For each defaulting firm, we draw a recovery rate from the state-dependent recovery rate distribution.

The basic idea

The DGP works as follows:

- The state of the credit cycle is determined by a two-state Markov chain.
- The number of defaulting firms is drawn using the state-dependent default probability.
- For each defaulting firm, we draw a recovery rate from the state-dependent recovery rate distribution.

Dependence:

- Conditional on the state, defaults are independent, recoveries between firms are independent, and the number of defaulting firms and recoveries are independent.
- As a consequence, (unconditional) dependence is driven entirely by the (unobserved) state of the credit cycle.

Specific assumptions about functional forms

Specific assumptions about functional forms

- Conditional on the state of the cycle, default arrival is described by discrete hazards of the form

$$\lambda_t = \left(1 + \exp\left\{\gamma_0 + \gamma_1 c_t + \gamma_2 X_t\right\}\right)^{-1}.$$

(t: time, c_t: cycle, X_t: economy-wide variables)

Specific assumptions about functional forms

- Conditional on the state of the cycle, default arrival is described by discrete hazards of the form

$$\lambda_t = \left(1 + \exp\left\{\gamma_0 + \gamma_1 c_t + \gamma_2 X_t\right\}\right)^{-1}.$$

  (t: time, c_t: cycle, X_t: economy-wide variables)
- Recoveries for each default event are drawn from a beta distribution.

Specific assumptions about functional forms

- Conditional on the state of the cycle, default arrival is described by discrete hazards of the form

$$\lambda_t = \left(1 + \exp\left\{\gamma_0 + \gamma_1 c_t + \gamma_2 X_t\right\}\right)^{-1}.$$

(t: time, c_t: cycle, X_t: economy-wide variables)

- Recoveries for each default event are drawn from a beta distribution.
- The parameters of this beta distribution are given by:

$$\alpha_{tis} = \exp\left\{\delta_0 + \delta_1 c_t + \cdots + \delta_6 X_t\right\} \quad (1)$$

$$\beta_{tis} = \exp\left\{\zeta_0 + \zeta_1 c_t + \cdots + \zeta_6 X_t\right\} \quad (2)$$

(i: firm, s: seniority)

Estimation

- The model can be easily be estimated using a version of the Hamilton filter (MLE).
- For this we need the number of defaulting firms, and non-defaulting firms in each period, and a recovery rate for each default event.

Data sources

We use data from two main sources,

Data sources

We use data from two main sources,

- the Altman-NYU Salomon Center Corporate Bond Default Master Database, consisting of prices just after default of more than 2,000 bonds of US firms from 1974 to 2005, with issuers and dates, and a "bond category",

Data sources

We use data from two main sources,

- the Altman-NYU Salomon Center Corporate Bond Default Master Database, consisting of prices just after default of more than 2,000 bonds of US firms from 1974 to 2005, with issuers and dates, and a “bond category”,
- and the annual default rates reported in Standard & Poor’s Quarterly Default Update from May 2006.

Data sources

We use data from two main sources,

- the Altman-NYU Salomon Center Corporate Bond Default Master Database, consisting of prices just after default of more than 2,000 bonds of US firms from 1974 to 2005, with issuers and dates, and a "bond category",
- and the annual default rates reported in Standard & Poor's Quarterly Default Update from May 2006.

Assuming that both the Altman data and Standard & Poor's data exhibit the same default rates, we can obtain the number of non-defaulting firms in each year.

Data sources

We use data from two main sources,

- the Altman-NYU Salomon Center Corporate Bond Default Master Database, consisting of prices just after default of more than 2,000 bonds of US firms from 1974 to 2005, with issuers and dates, and a "bond category",
- and the annual default rates reported in Standard & Poor's Quarterly Default Update from May 2006.

Assuming that both the Altman data and Standard & Poor's data exhibit the same default rates, we can obtain the number of non-defaulting firms in each year.

- We augment this with GDP growth, investment growth, unemployment, the S&P 500 index, the VIX, the slope of the term structure, and an NBER recession indicator.

Macro variables versus the business cycle

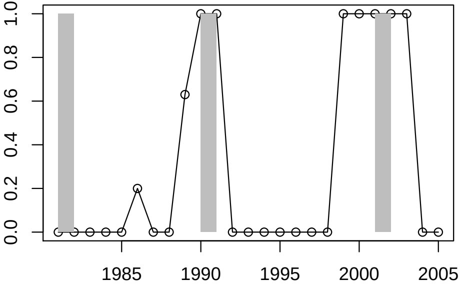
	Model 1	Model 2	Model 3	Model 4
Explanatory variables for λ (default probability)				
	constant	constant	constant	constant
		cycle	log GDP growth	cycle
				log GDP growth
Explanatory variables for α,β (recovery rates)				
	constant	constant	constant	constant
		cycle	log GDP growth	cycle
				log GDP growth
	seniority[a]	seniority[a]	seniority[a]	seniority[a]
AIC	-37.73	-405.31	-128.96	-412.79
BIC	0.0727	-0.1549	0.0188	-0.1290

Macro variables versus the business cycle

	Model 1	Model 2	Model 3	Model 4
Explanatory variables for λ (default probability)				
	constant	constant	constant	constant
		cycle	log GDP growth	cycle
				log GDP growth
Explanatory variables for α,β (recovery rates)				
	constant	constant	constant	constant
		cycle	log GDP growth	cycle
				log GDP growth
	seniority[a]	seniority[a]	seniority[a]	seniority[a]
AIC	-37.73	-405.31	-128.96	-412.79
BIC	0.0727	-0.1549	0.0188	-0.1290

$\Longrightarrow$ Macro variables are significant, but do not contribute much to the fit.

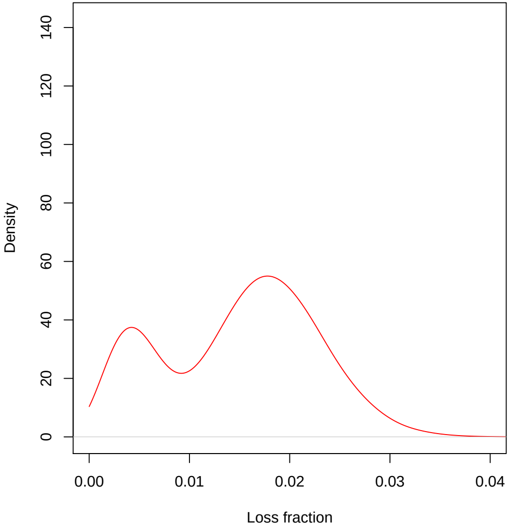
Is the business cycle = credit cycle?



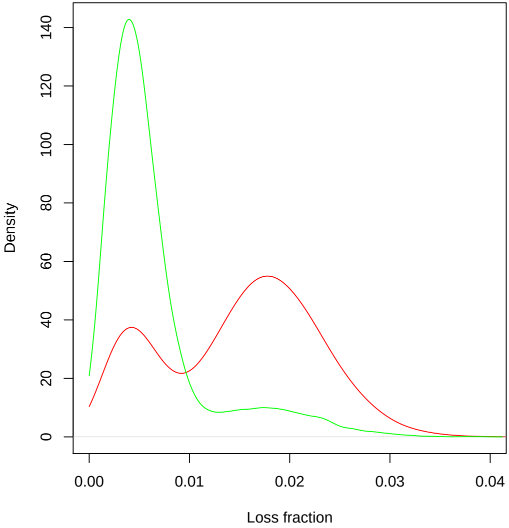
- The estimated credit downturns start earlier than NBER recessions, and end later.
- We investigate lead-lag relationships between macro variables and credit variables and find that recovery rates Granger cause log GDP growth (very significant!).

VaR Simulation (1)

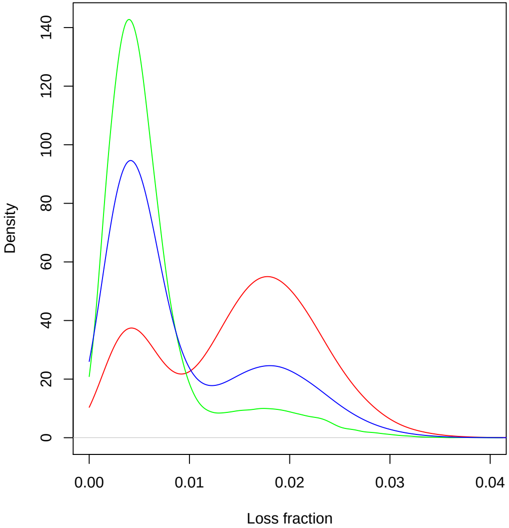
VaR Simulation (1)



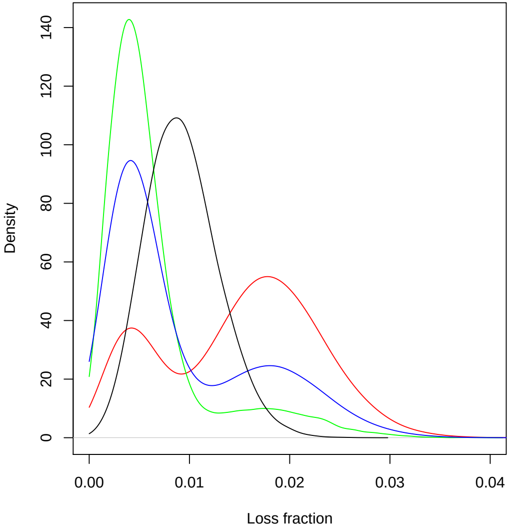
VaR Simulation (1)



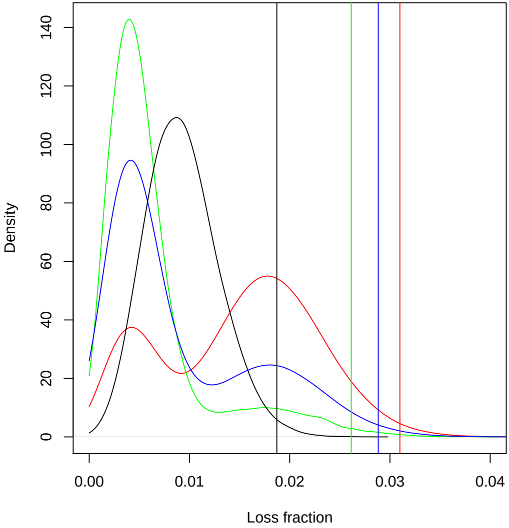
VaR Simulation (1)



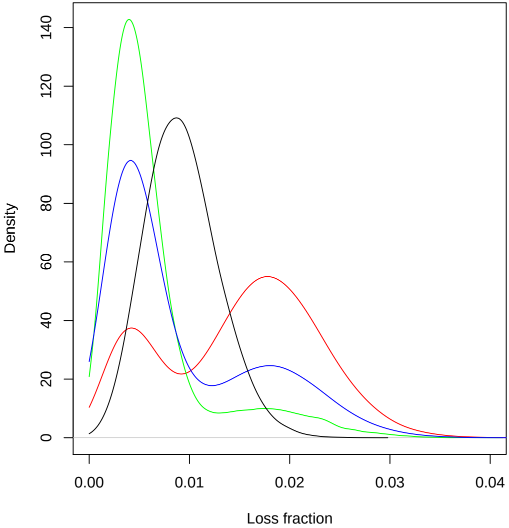
VaR Simulation (1)



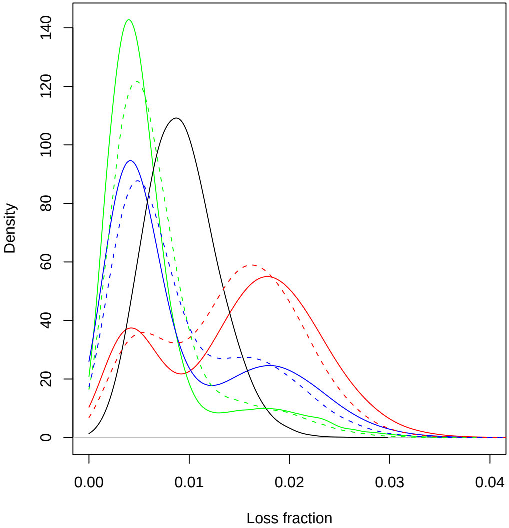
VaR Simulation (1)



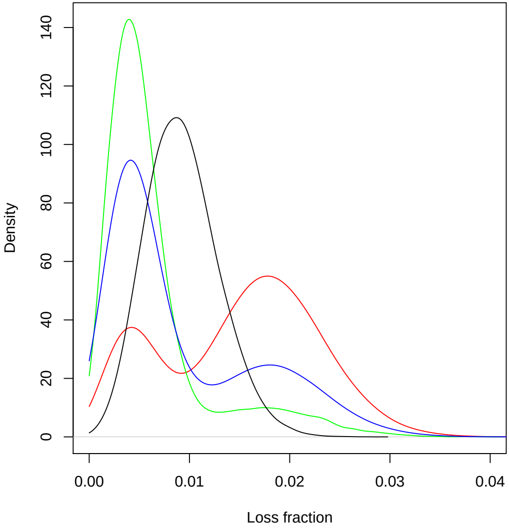
VaR Simulation (2)



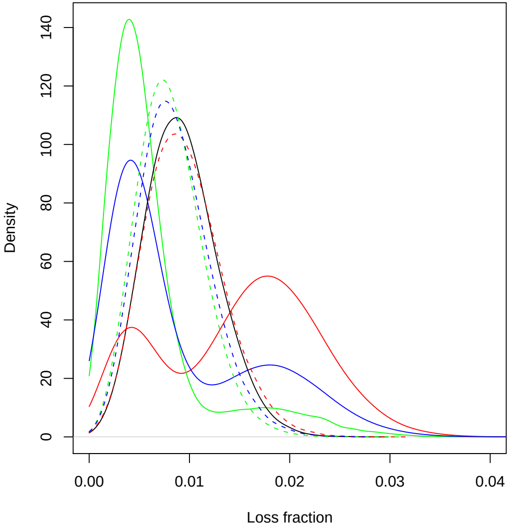
VaR Simulation (2)



VaR Simulation (3)



VaR Simulation (3)



Expected Loss

Expected loss is not affected

- Suppose

$$PD|E = 2\% \quad L|E = 30\%$$
$$PD|R = 10\% \quad L|R = 70\%$$
$$P(E) = 1/2$$

Expected Loss

Expected loss is not affected

- Suppose

$$
\begin{array}{rl}
PD|E = 2\% & L|E = 30\% \\
PD|R = 10\% & L|R = 70\%
\end{array}
$$

$$P(E) = 1/2 \implies E[L] = 50\%, E[PD] = 6\%.$$

Expected Loss

Expected loss is not affected

- Suppose

$$PD|E = 2\% \quad L|E = 30\%$$
$$PD|R = 10\% \quad L|R = 70\%$$
$$P(E) = 1/2 \implies E[L] = 50\%, E[PD] = 6\%.$$

Therefore

$$E[L \cdot PD] = .5 \times 30\% \times 0.02 + .5 \times 70\% \times 0.1 = 3.8\%$$
$$E[L] \cdot E[PD] = 3\%$$

$$\implies E[L \cdot PD] - E[L] \cdot E[PD] = 80bp$$

Expected Loss

Expected loss is not affected

- Suppose
  $$PD|E = 2\% \quad L|E = 30\%$$
  $$PD|R = 10\% \quad L|R = 70\%$$
  $P(E) = 1/2 \implies E[L] = 50\%, E[PD] = 6\%.$
  Therefore
  $$E[L \cdot PD] = .5 \times 30\% \times 0.02 + .5 \times 70\% \times 0.1 = 3.8\%$$
  $$E[L] \cdot E[PD] = 3\%$$
  $$\implies E[L \cdot PD] - E[L] \cdot E[PD] = 80bp$$
- But $E[L \cdot PD] - E[L]E[PD] = \mathrm{Cov}(L, PD)$.

Expected Loss

Expected loss is not affected

- Suppose
  $PD|E = 2\% \quad L|E = 30\%$
  $PD|R = 10\% \quad L|R = 70\%$
  $P(E) = 1/2 \implies E[L] = 50\%, E[PD] = 6\%.$
  Therefore

$$E[L \cdot PD] = .5 \times 30\% \times 0.02 + .5 \times 70\% \times 0.1 = 3.8\%$$
$$E[L] \cdot E[PD] = 3\%$$

$$\implies E[L \cdot PD] - E[L] \cdot E[PD] = 80bp$$

- But $E[L \cdot PD] - E[L]E[PD] = \mathrm{Cov}(L, PD)$.
- In our data, $\mathrm{Cov}(avg.\, L, dfr) = 5bp$.

Some conclusions

- We propose an econometric model in which default rates and recovery rates are driven by an unobserved Markov chain.
- This describes the data well, and does better than many models based on observed covariates.

We can use the estimated model to look at what happens when we go from constant to time-varying recovery rate distributions. We get

- higher estimates of tail risk,
- the same expected losses,
- bigger swings in spreads over the cycle. Average spreads over the cycle are not affected.