BANCO DE ESPAÑA
Eurosistema

GABRIEL JIMÉNEZ AND JAVIER MENCÍA

Modelling the distribution of credit losses with observable and latent factors

Javier Mencía

Third International Conference on Credit and Operational Risks

Montreal
12-13 April 2007

FINANCIAL STABILITY

- Credit risk is one of the variables more directly related to financial stability.
- Basel II has put forward the need to measure this risk accurately.
- As a consequence, several models have been proposed to assess credit risk from a systemic point of view:
  - Austria: Boss (2002)
  - Finland: Virolainen (2004)
  - U.K. : Drehmann (2005) y Drehmann, Patton y Sorensen (2006)
  - International: Pesaran, Schuermann, Treutler y Weiner (2006)

- **The main characteristics of these papers are:**
  - –Analysis of Credit risk across different sectors,
  - –Effect of macroeconomic variables on default frequencies.

- **However:**
  - –They do not allow for contagion effects due to unobservable or not modelled factors,
  - –They do not model the growth of the loan market size,
  - –They do not consider loans to individuals, such as mortgages or consumption loans.

Contributions of this paper

- We develop a model to estimate the credit loss distribution of the loans in a banking system.
- There are 10 corporate sectors plus mortgages and consumption loans.
- We consider the effect of macro variables (GDP, interest rates, ...) and allow for contagion effects through latent factors.
- We apply our model to the Spanish credit market, where we also carry out stress tests.

Plan of the presentation

- **Introduction**
- **Theoretical part: model and simulation.**
- **Empirical application: data, model specifications and stress tests.**
- **Conclusions.**

Model

- Consider an economy with *K* sectors. We will express the losses due to loan *i* at time *t* as:

$$L_{i,k,t} = D_{i,k,t} LGD_{k,t} EAD_{i,k,t}$$

where

- $D_{i,k,t}$=1 if i defaults and 0 otherwise
- $EAD_{i,k,t}$: Exposure at default
- $LGD_{k,t}$ in (0,1): Loss Given default

Model

- **Total losses in sector k:**

$$L_{k,t} = \sum_{i=1}^{n_{k,t}} D_{i,k,t} LGD_{k,t} EAD_{i,k,t}$$
$$= LGD_{k,t} \underbrace{\sum_{i=1}^{n_{k,t} \cdot p_{k,t}} EAD_{i,k,t}}_{S_k(p_{kt} \cdot n_{kt})},$$

where

–n_{kt}: total number of loans of sector k.

–p_{kt}: default frequency, i.e. number of non performing loans divided by n_{kt}.

Model

- **Hence, our model consists of the following stochastic components:**
  - $n_{k,t}$, p_{kt}: joint dynamic model
  - $EAD_{i,k,t}$: Fit a positive distribution (Inverse Gaussian, Beta)
  - $LGD_{i,k}$: Beta distribution

Evolution of n_{kt} and p_{kt}

We consider the Gaussian VAR model:

$$\Delta n_{k,t} = \alpha_{1,k} + \sum_{j=1}^{q} \rho_{1,j} \Delta n_{k,t-j} + \sum_{j=1}^{r} \boldsymbol{\gamma}'_{1,j} \mathbf{x}_{t-j} + \boldsymbol{\beta}_{1,k} f_{1,t} + u_{1,kt},$$

$$y_{k,t} = \sum_{j=1}^{q} \rho_{2,j} y_{k,t-j} + \sum_{j=1}^{r} \boldsymbol{\gamma}'_{2,j} \mathbf{x}_{t-j} + \boldsymbol{\beta}_{2,k} f_{2,t} + u_{2,kt},$$

where

$$\Delta n_{k,t} = \log(n_{k,t}) - \log(n_{k,t-1}),$$

$$y_{k,t} = \Phi^{-1}(p_{k,t}) - \Phi^{-1}(p_{k,t-1}),$$

$f_t=(f_{1t},f_{2t})'$ are latent factors independent of m macro variables x_t, with

$$\mathbf{x}_t = \delta_0 + \sum_{j=1}^{s} \mathbf{A}_j \mathbf{x}_{t-j} + \mathbf{v}_t$$

$$\mathbf{f}_t = \mathbf{R}\mathbf{f}_{t-1} + \mathbf{w}_t$$

Exposures at default

Because of their properties, we consider these two distributions to fit EAD data:

- **Inverse Gaussian: IG(μ_k,λ_k)**

$$f_{IG}\left(EAD_{i,k,t} = x; \mu_k, \lambda_k\right) = \left(\frac{\lambda_k}{2\pi x^3}\right) \exp\left[-\frac{\lambda_k}{2\mu_k^2 x}\left(x - \mu_k\right)^2\right]$$

where the mean is μ_k and the variance is μ_k^3/λ_k

- **Gamma: Gamma(ν_k,τ_k)**

$$f_{Gamma}(EAD_{i,k,t} = x; \nu_k, \tau_k) = \frac{(x/\tau)^{\nu/2-1}}{2^{\nu/2}\Gamma(\nu/2)\tau} \exp\left(\frac{-\nu x}{2\tau}\right)$$

where the mean is $\nu_k\tau_k$ and the variance is $\nu_k\tau_k^2$

Exposures at default

- We exploit the fact that both the IG and the Gamma are closed under aggregation.

- Specifically, the distribution of

$$S_k(p_{kt} \cdot n_{kt}) = \sum_{i=1}^{n_{k,t} \cdot p_{k,t}} EAD_{i,k,t}$$

will be either IG($p_{kt}n_{kt}\mu_k$,$(p_{kt}n_{kt})^2\lambda_k$) or Gamma($p_{kt}n_{kt}\nu_k$,$\tau_k$).

- We also analyse the effect of the macroeconomic shocks on the means of the EAD's with the following dynamic parametrisation of the mean:

$$\mu_{kt} = \mu_{kt-1} \exp\left[\boldsymbol{\varphi}_k' \mathbf{v}_{t-1} - (1/2)\boldsymbol{\varphi}_k' \boldsymbol{\Omega} \boldsymbol{\varphi}_k\right]$$

where v_t=x_t-E_{t-1}(x_t) and Ω=V(v_t).

Estimation of the credit loss distribution

We estimate this distribution by simulation, repeating the following steps:

1. **Generate a random draw of the vector of macro variables x_t.**
2. **Generate n_{kt} and p_{kt} from the VAR $\Rightarrow$ Number of defaults: $n_{kt} \cdot p_{kt}$**
3. **Generate the sum of the exposures at default, $S_k(p_{kt} \cdot n_{kt})$, from either the IG or the Gamma distributions.**
4. **Generate $LGD_{k,t}$ from the Beta distribution.**
5. **Total losses in sector k:**

$$L_{k,t} = LGD_{k,t} S_k(p_{kt} \cdot n_{kt}), \quad k = 1, \cdots, K$$

Empirical application to the Spanish banking system

- We use the Spanish credit register, which reports data of every loan with an exposure above € 6000.

- We have obtained Quarterly series of n_{kt}, p_{kt} and EAD_{ikt} from 1984Q4 to 2006Q3.
- Unfortunately, LGD data is not available.

- We classify loans in the following groups:
  - Corporate sectors: (1) Agriculture, (2) Mining, (3) Manufacture, (4) Utilities, (5) Construction, (6) Commerce, (7) Hotels, (8) Communications, (9) R&D and (10) Other Corporate

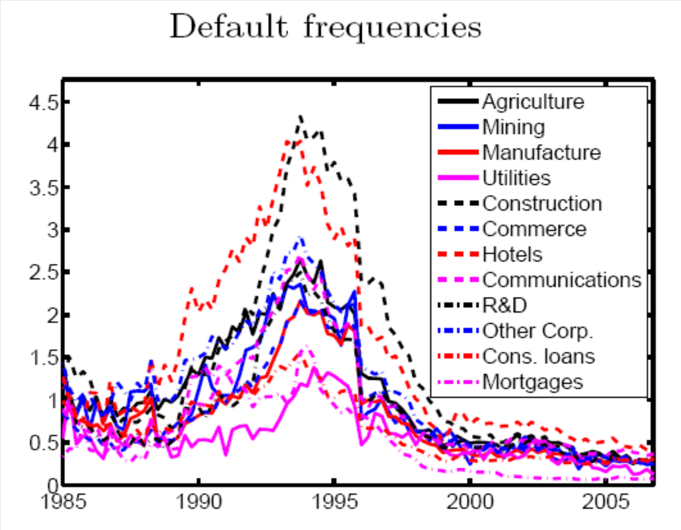
  - Individuals: (11) Consumption loans and (12) Mortgages

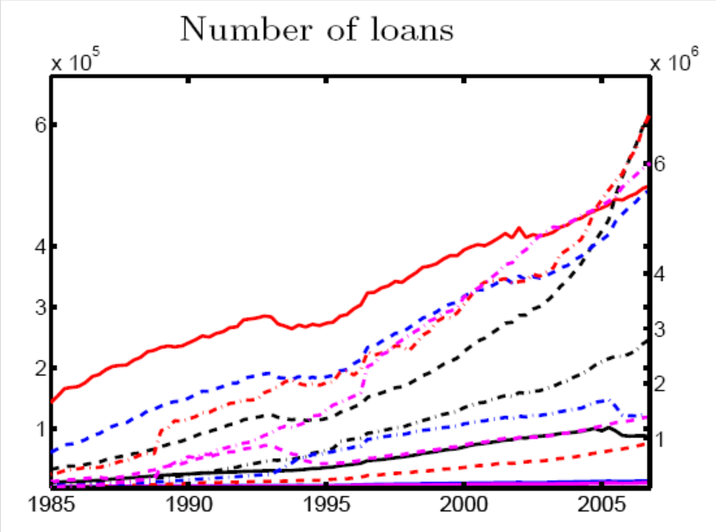
- **Since we do not have LGD data, we choose the mean of the betas with the results reported in the QIS5, while the standard deviations are fixed to 20%.**

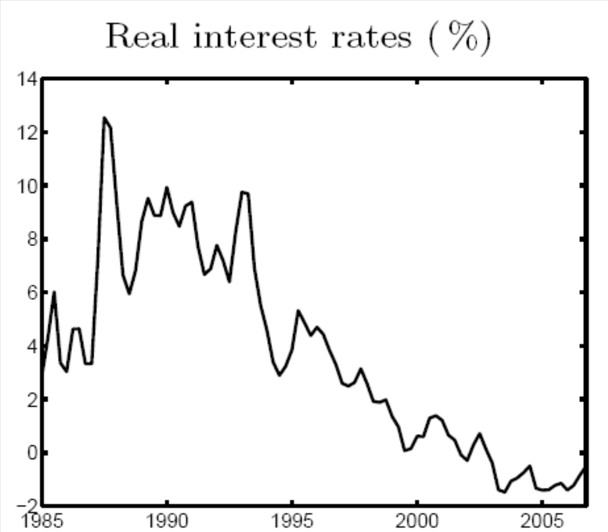
- **We compare three different specifications of our model:**
  1. **GDP and (real) interest rates, but not latent factors,**
  2. **GDP, interest rates and latent factors,**
  3. **GDP, interest rates, spread, latent factors, unemployment, and production series by corporate sector used as sectorial characteristics.**

- **In addition, we estimate the credit loss distribution with a static and a dynamic model for exposures at default.**

Historical evolution of the series







Model of the default frequencies (p_{kt})

	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	f_{2t}
Agriculture	-1.245**	-1.124**	-0.259	-0.238	1.527**	-0.317	3.356**
Mining	-1.556	-1.416	0.238	0.310	0.423	-1.019	5.781**
Manufacture	-1.692**	-1.979**	-0.838	0.426	0.731	-0.392	4.434**
Utilities	-0.338	-0.410	-0.713	0.031	0.724	-0.790	5.071**
Construction	-1.133**	-1.098**	-0.828*	0.703	0.163	0.293	3.448**
Commerce	-1.404**	-1.322**	-0.514	-0.175	0.775	-0.081	4.050**
Hotels	-1.408**	-1.027	-0.207	-0.103	1.879**	-0.309	4.059**
Communications	-1.156**	-1.247**	-0.859*	0.129	1.195**	-0.401	3.656**
R&D	-0.566	-1.637**	-1.478**	0.142	-0.105	-0.059	3.736**
Other Corp.	-0.447	-1.013*	-0.253	0.682	0.913*	-0.209	3.173**
Consumption loans	-0.897**	-1.129**	-0.547	0.053	0.630	0.293	3.353**
Mortgages	-0.957	-1.910**	-1.342**	0.367	0.097	0.154	1.644**

Model of the evolution of the number of loans (n_{kt})

	GDP_{t-2}	GDP_{t-3}	GDP_{t-4}	INT_{t-2}	INT_{t-3}	INT_{t-4}	f_{1t}
Agriculture	0.274	0.234	0.151	-0.192	0.048	-0.088	1.257**
Mining	0.206	-0.152	0.021	-0.046	-0.064	0.215	1.391**
Manufacture	0.409**	0.187	0.097	-0.070	-0.078	0.068	1.605**
Utilities	0.279	-0.047	-0.084	-0.866**	0.560	-0.505	1.217**
Construction	0.337*	0.151	0.106	-0.242	0.060	-0.136	1.487**
Commerce	0.483**	0.213	0.043	0.083	-0.206	0.140	1.800**
Hotels	0.214	0.003	0.071	0.031	0.029	-0.251	2.019**
Communications	0.122	0.643	0.430	0.629	-0.115	0.121	2.066**
R&D	0.675**	0.356	-0.052	-0.064	-0.094	-0.233	1.569**
Other Corp.	-0.998**	-0.786*	0.236	0.406	-0.239	0.549	1.068**
Consumption loans	-0.039	0.065	0.544*	0.504	0.313	0.035	0.849**
Mortgages	0.134	0.054	0.134	0.763**	-0.510	-0.129	0.606*

Dynamics of the explanatory variables

	Intercept	Lag 1	Lag 2	Conditional covariance matrix			
				GDP	INT	f_{1t}	f_{2t}
GDP	0.034	-0.435**	-0.040	1.187**			
INT	-0.118	0.550**	-0.513**	-0.101	0.955**		
f_{1t}	0	-0.195*	0	0	0	1	
f_{2t}	0	0.231**	0	0	0	-0.473**	1

Dynamic model of the means of EAD's

	Mean in 2006.Q3	η_k	GDP_{t-1}	INT_{t-1}
Agriculture	0.103	-0.003**	-0.038**	0.159**
Mining	0.087	-0.020**	0.003	0.057
Manufacture	0.089	-0.014**	-0.019**	0.028**
Utilities	0.188	0.026**	-0.151**	-0.179**
Construction	0.089	-0.024**	-0.062**	0.067**
Commerce	0.086	-0.009**	-0.033**	0.029**
Hotels	0.065	-0.025**	-0.090**	0.042**
Communications	0.055	-0.018**	-0.051**	0.008*
R&D	0.055	-0.016**	-0.103**	0.026**
Other Corp.	0.099	-0.014**	-0.021**	0.029**
Consumption loans	0.016	-0.019**	0.025**	0.018**
Mortgages	0.058	0.003**	-0.043**	0.018**

Note: We estimate $\mu_{kt} = \mu_{kt-1} \exp\left[\eta_k + \boldsymbol{\varphi}_k' \mathbf{v}_{t-1} - (1/2)\boldsymbol{\varphi}_k' \boldsymbol{\Omega} \boldsymbol{\varphi}_k\right]$, where η_k captures a time trend.

Fit of empirical correlations

- **We test whether latent factors are able to capture the intersectorial correlations of default frequencies and the evolution of the number of loans.**
- **For example, the conditional correlation between (transformed) default frequencies can be expressed as:**

$$\mathrm{corr}_{t-1}(y_{it}, y_{jt}) = \beta_{2,i}\beta_{2,j}$$

- **This correlation will be zero in the absence of latent factors.**

- **We test this hypothesis by computing the correlations between the fitted residuals:**

$$\varepsilon_{kt}(\hat{\theta}) = y_{kt} - E_{t-1}(y_{kt}; \hat{\theta})$$

Fit of empirical correlations (p-values)

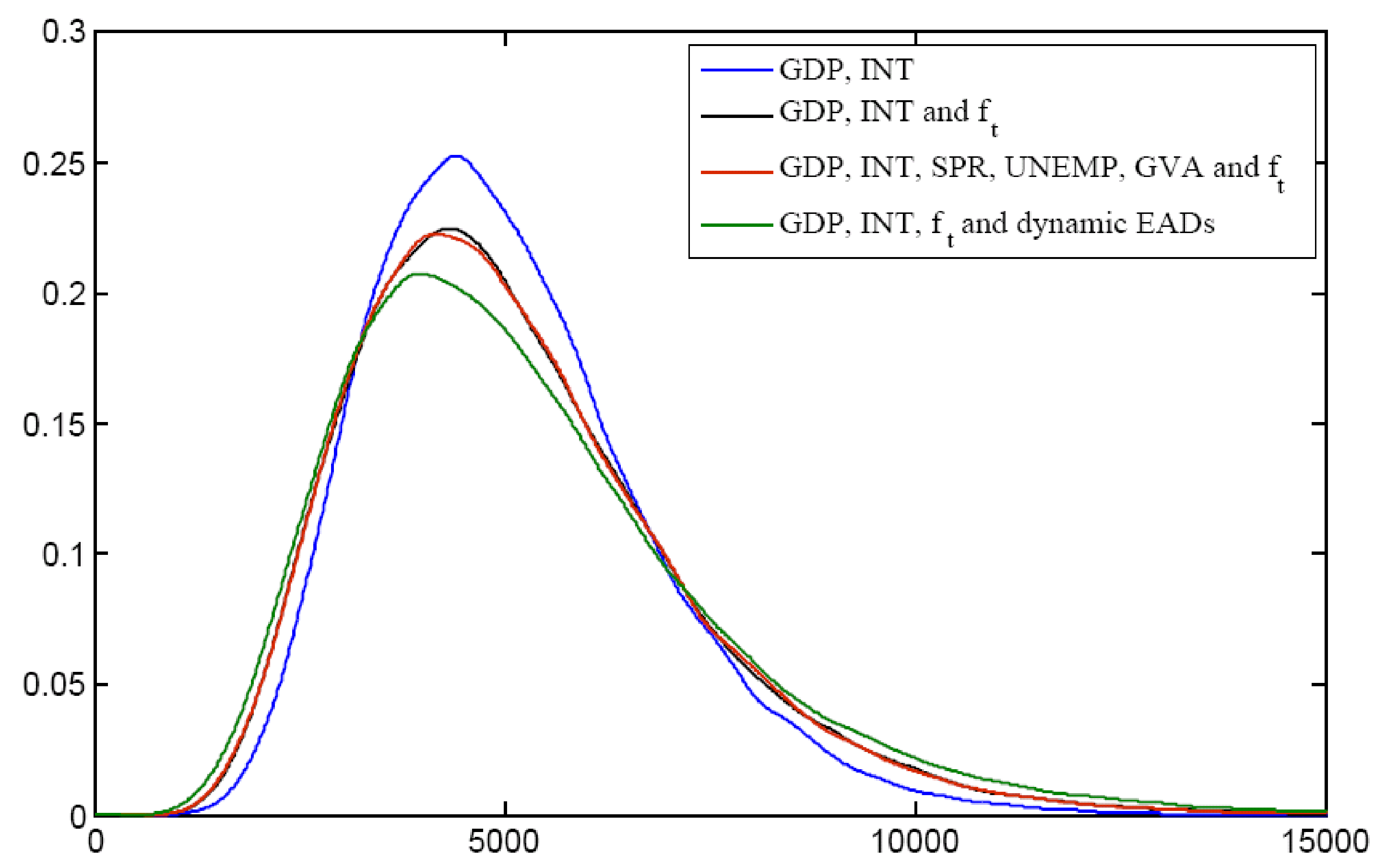
(a) Model with GDP and Interest rates

		1	2	3	4	5	6	7	8	9	10	11
Agriculture	1											
Mining	2	0.00										
Manufacture	3	0.00	0.00									
Utilities	4	0.00	0.00	0.00								
Construction	5	0.00	0.00	0.00	0.00							
Commerce	6	0.00	0.00	0.00	0.00	0.00						
Hotels	7	0.00	0.00	0.00	0.00	0.00	0.00					
Communications	8	0.00	0.00	0.00	0.00	0.00	0.00	0.00				
R&D	9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00			
Other Corp.	10	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		
Cons. loans	11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
Mortgages	12	0.03	0.16	0.06	0.06	0.01	0.11	0.27	0.10	0.01	0.00	0.00

(b) Model with GDP, Interest rates and latent factors

		1	2	3	4	5	6	7	8	9	10	11
Agriculture	1											
Mining	2	0.32										
Manufacture	3	0.68	0.03									
Utilities	4	0.80	0.70	0.94								
Construction	5	0.75	0.23	0.57	0.15							
Commerce	6	0.68	0.98	0.70	0.62	0.37						
Hotels	7	0.43	0.27	0.50	0.39	0.73	0.98					
Communications	8	0.69	0.55	0.86	0.96	0.46	1.00	0.75				
R&D	9	0.41	0.14	0.13	0.10	0.43	0.33	0.98	0.48			
Other Corp.	10	0.71	0.66	0.52	0.15	0.31	0.37	0.72	0.31	0.01		
Cons. loans	11	0.36	0.30	0.17	0.99	0.45	0.24	0.71	0.21	0.29	0.83	
Mortgages	12	0.79	0.68	0.56	0.55	0.47	0.46	0.55	0.57	0.42	0.41	0.19

Credit loss distributions at a three year horizon for (2006Q3)



Kernel based densities obtained from 100,000 simulations

Descriptive statistics of the loss distribution (€ Million)

	Expected loss			VaR(99,9 %)		
	1 year	3 years	5 years	1 year	3 years	5 years
Agriculture	42.01	146.15	283.52	159.10	961.08	2822.62
Mining	4.78	16.43	31.40	21.91	103.52	254.48
Manufacture	204.64	662.15	1194.00	692.44	2974.53	6647.74
Utilities	4.30	14.64	28.26	27.82	157.63	469.12
Construction	305.93	1121.13	2271.62	1094.45	6536.97	18775.45
Commerce	187.63	629.52	1175.78	620.89	2789.91	6461.51
Hostels	34.65	124.20	250.97	130.64	717.79	1993.80
Communications	35.72	121.15	230.34	121.04	595.81	1506.49
R&D	67.90	256.04	545.97	252.31	1604.70	4998.29
Other Corp.	54.92	195.07	388.18	182.99	952.80	2586.82
Consumption loans	383.46	1290.97	2420.16	1592.82	6360.12	13727.51
Mortgages	182.43	696.99	1521.42	1116.57	6926.87	21625.75
Total	1508.37	5274.44	10341.61	3507.63	17499.98	47652.85

We stress the credit loss distribution at 2006Q3 with artificial shocks of 3 s.d. that occur in the first quarter of our three-year horizon.

Change caused by the shocks (%)

	- 3 s.d. GDP shock		+ 3 s.d. INT shock	
Model specification	ES	VaR (99.9 %)	ES	VaR (99.9 %)
GDP,INT,f_t	17	16	7	8
GDP,INT,SPR,UNEMP,GVA,f_t	16	15	7	8
GDP,INT,f_t and dynamic EAD's	30	31	18	18

Conclusions

- This paper develops a flexible model to estimate the credit loss distribution of loans in a banking system.

- We analyse the impact macroeconomic events on default frequencies, the total number of loans and the distribution of exposures at default.

- We also allow for contagion effects due to unobservable factors.

- We estimate the credit loss distribution by simulation from our model through a computationally fast and efficient methodology.

Conclusions

- We consider an application to the Spanish banking system.
- 10 corporate sectors and 2 groups of loans to individuals.
- Our results show a strong dependence of Spanish loans on macroeconomic characteristics, specially GDP.
- Latent factors are also highly significant and cause fatter tails in the credit loss distribution.
- In absolute terms, construction, manufacture, consumption loans and mortgages are the groups of higher risk.
- Finally, we perform stress tests that show a higher sensitivity to GDP shocks in our application.

GABRIEL JIMÉNEZ AND JAVIER MENCÍA

THANKS FOR YOUR ATTENTION

BANCO DE ESPAÑA
Eurosistema