

MODELLING WEEK 2013

Impact of the real estate boom & crash on the mortgages recovery indicators



Instructores:

**Alejandro González Salcedo
Ignacio Villanueva**

**Patricia Blázquez Esquinas
Mark Gilbert
Marta López Caamaño
Nair Quintas Rodríguez
Víctor Darío Rodríguez Rodríguez
Leonardo Torres Hansa**

14 de junio de 2013

Phases of the problem:

1. Calculate LGD (Loss given default) with buckets of LTV (loan to value)
2. Search housing prices
3. Calculate associated capital requirements for the financial institution

We suppose we are the bank.

Description of the variables:

Nombre variable	Descripción
ID_contrato	Identificador único del contrato
F_MOVIMIENTO	Fecha en que se recoge el movimiento contable
IMPORTE_MOVIMIENTO	Importe debido al movimiento
TIPO_MOVIMIENTO	Tipo de movimiento por el que se ha efectuado el movimiento
F_EVENTO	Fecha de entrada en default del proceso recuperatorio
F_APERTURA	Fecha de formalización del contrato
VALOR_TASACION	Valor de la tasación a fecha de formalización del contrato
EXPOSICION_EAD	Exposición en el momento del default
F_SALIDA_EVENTO	Fecha de salida del default. Fin del periodo recuperatorio
FINALIZACION	Tipo de finalización del contrato: ABIERTA, CURADA, ADJUDICACION
COSTES_INDIRECTOS	Costes indirectos debidos a los movimientos contables del contrato

ID_contrato: contract ID

F_Movimiento: Movement date

Importe_Movimiento: Amount transfered

Tipo_Movimiento: Type of movement. There are three types of movement:

- **Recobro:** Payments made by the client (positive for the bank).
- **Gastos:** Payments made by the bank. These include lawyer expenses, technical costs...
- **Adjudicación:** Foreclosure. It occurs when the bank takes the collateral of the loan. In this case, the collateral is the house so there's a positive movement for the bank equivalent to the value of the property.

F_evento: Default date. A contract goes to default when 90 days pass without paying.

F_apertura: Start date of the contract.

Valor_tasación: Independently evaluated value of property at F_apertura.

Exposición_EAD: Exposure at Default, amount of money owed to bank at default date.

F_Salida_Evento: Date of resolution.

Finalización: Type of resolution. A contract can result into:

- **Adjudicación:** Foreclosure, the bank takes the house.
- **Curada:** Cured, the client has brought his payments up to date.

- **Fallido:** Failed, the client has made no payments for at least four years so the operation is closed and the bank absorbs the debt. In this case the bank will need an amount of capital to cover the possible costs.

Costes_Indirectos: Indirect costs to the bank. It is a weighted average of the general cost of the bank (offices, bank employees ...)

Data

We have several data sets. The first includes the payments for each contract, and the date and type of payment.

	A	B	C	D	E
		ID_contrato	F_MOVIMIENTO	IMPORTE_MOVIMIENTO	TIPO_MOVIMIENTO
4					
5		1	30/07/2002	49.382,39	ADJUDICACION
6		1	26/11/2001	-66,11	GASTO
7		1	20/03/2003	-1.205,77	GASTO
8		1	02/11/2000	1.060,15	RECOBRO
9		1	30/07/2002	4.111,77	RECOBRO
10		1	20/03/2003	1.205,77	RECOBRO
11		2	30/08/2001	10.535,95	RECOBRO
12		2	19/01/2002	18.322,05	RECOBRO
13		3	01/06/2004	2.384,48	RECOBRO
14		3	13/01/2005	3.327,64	RECOBRO
15		3	10/01/2006	2.484,15	RECOBRO
16		4	27/11/2002	1.786,90	ADJUDICACION
17		4	25/10/2002	583,71	RECOBRO
18		5	24/01/2004	16.954,32	ADJUDICACION
19		5	12/02/2002	-190,39	GASTO
20		5	01/07/2004	-2.696,54	GASTO
21		5	24/01/2004	48.707,30	RECOBRO
22		6	01/07/2005	3.003,66	RECOBRO
23		6	07/04/2006	4.894,00	RECOBRO
24		6	03/05/2007	4.732,79	RECOBRO
25		6	11/04/2008	4.593,05	RECOBRO
26		6	13/05/2009	4.442,24	RECOBRO
27		6	09/04/2010	4.284,23	RECOBRO
28		6	10/05/2011	4.138,95	RECOBRO
29		6	22/05/2012	2.358,22	RECOBRO
30		7	03/01/2001	66.365,45	RECOBRO

The second data set corresponds to closed contracts. This is the historical data we will analyze and use to forecast the open contracts. For this data we know:

	A	B	C	D	E	F	G	H	I
1	TABLA DE PARTIDA CON LA INFORMACIÓN ASOCIADA A CADA PERIODO RECUPERATORIO CERRADO								
2									
3									
4									
5		ID_CONTRATO	F_EVENTO	F_APERTURA	VALOR_TASACION	EXPOSICION_EAD	F_SALIDA_EVENTO	FINALIZACION	COSTES_INDIRECTOS
6		1	03/10/2000	12/01/2000	52.345,48	47.269,54	20/03/2003	ADJUDICACION	0,427
7		2	04/07/2001	14/01/2000	227.993,44	134.577,72	30/08/2001	CURADA	0,005
8		3	04/04/2004	04/01/2000	47.099,28	37.996,30	26/07/2004	CURADA	0,011
9		4	04/07/2002	12/01/2000	23.008,00	29.258,52	27/11/2002	ADJUDICACION	0,366
10		5	02/11/2001	03/03/2000	61.805,43	54.744,09	01/07/2004	ADJUDICACION	0,389
11		6	23/06/2005	20/01/2000	90.559,84	33.419,77	01/07/2005	CURADA	0,001
12		7	03/10/2000	17/01/2000	151.455,05	63.766,90	31/05/2003	FALLIDO	0,427
13		8	03/02/2009	14/01/2000	48.994,51	24.231,42	20/02/2009	CURADA	0,002
14		9	03/09/2002	21/02/2000	167.790,56	62.115,67	15/11/2002	CURADA	0,007
15		10	05/04/2006	09/02/2000	14.070,11	10.600,13	29/04/2010	ADJUDICACION	0,234
16		11	03/02/2001	28/01/2000	21.636,44	14.927,68	23/03/2001	CURADA	0,005
17		12	03/09/2000	24/01/2000	81.723,07	25.790,96	30/05/2004	FALLIDO	0,430
18		13	03/02/2004	17/02/2000	43.423,12	32.711,52	10/05/2006	FALLIDO	0,310
19		14	25/12/2000	28/01/2000	293.293,90	257.553,20	03/03/2001	CURADA	0,007
20		15	03/11/2000	31/01/2000	73.599,83	103.097,66	01/02/2001	ADJUDICACION	0,424
21		16	03/09/2010	08/02/2000	42.854,17	32.147,80	07/12/2011	ADJUDICACION	0,079
22		17	23/04/2003	22/02/2000	167.409,34	33.953,33	13/09/2009	FALLIDO	0,338
23		18	03/09/2004	10/02/2000	60.101,21	31.840,29	25/02/2005	CURADA	0,017
24		19	03/10/2001	07/02/2000	48.829,92	61.339,33	19/01/2004	ADJUDICACION	0,392
25		20	04/07/2000	07/02/2000	42.965,50	34.482,93	01/02/2005	ADJUDICACION	0,436
26		21	03/09/2001	11/02/2000	120.202,51	88.983,15	17/12/2009	FALLIDO	0,395
27		22	04/07/2001	15/02/2000	38.224,00	29.658,08	26/02/2005	FALLIDO	0,401
28		23	03/10/2001	09/02/2000	44.353,52	53.671,98	19/01/2004	ADJUDICACION	0,392
29		24	03/08/2002	29/03/2000	61.457,50	36.713,61	24/04/2003	CURADA	0,025
30		25	10/04/2002	08/03/2000	72.121,45	17.231,08	18/06/2002	CURADA	0,007
31		26	04/07/2000	09/03/2000	43.332,97	13.127,17	29/10/2007	FALLIDO	0,436
		27	04/07/2001	10/03/2000	123.236,16	98.368,68	15/05/2002	CURADA	0,030

The third data set corresponds to the open contracts. These are contracts that are still in default so we want to forecast the amount of money that the bank will end up losing. For these contracts we know the default date, the date when the contract was signed, the property value when the contract was opened, the exposure at default and the indirect costs:

TABLA DE PARTIDA CON LA INFORMACIÓN ASOCIADA A CADA PERIODO RECUPERATORIO ABIERTO

ID_CONTRATO	F_EVENTO	F_APERTURA(CONTRATO)	VALOR_TASACION	EXPOSICION_EAD	FINALIZACION	COSTES_INDIRECTOS
1_A	05/04/2009	16/12/1999	53.309,77	29.284,31	ABIERTA	0,128
2_A	13/01/2012	18/07/2001	213.117,93	131.449,79	ABIERTA	0,031
3_A	31/08/2009	04/07/2001	51.646,17	25.900,81	ABIERTA	0,114
4_A	05/09/2011	07/11/2001	75.272,43	46.819,42	ABIERTA	0,043
5_A	05/03/2012	28/11/2001	96.697,78	56.908,03	ABIERTA	0,026
6_A	03/02/2009	03/01/2002	93.177,31	51.761,40	ABIERTA	0,134
7_A	25/05/2010	06/02/2002	42.900,00	28.830,73	ABIERTA	0,089
8_A	03/08/2010	04/03/2002	75.626,33	70.432,39	ABIERTA	0,082
9_A	04/07/2009	27/02/2002	111.155,00	88.613,30	ABIERTA	0,120
10_A	03/11/2011	07/05/2002	123.750,16	77.599,85	ABIERTA	0,038
11_A	05/03/2012	14/05/2002	36.000,00	21.845,91	ABIERTA	0,026
12_A	05/12/2011	12/07/2002	50.833,84	29.197,99	ABIERTA	0,035
13_A	05/03/2009	28/08/2002	45.264,02	42.145,51	ABIERTA	0,131
14_A	03/04/2010	02/10/2002	71.089,00	37.656,43	ABIERTA	0,094
15_A	06/05/2011	04/10/2002	96.284,76	76.041,81	ABIERTA	0,055
16_A	23/12/2009	23/10/2002	74.600,00	50.808,19	ABIERTA	0,103
17_A	03/08/2011	29/11/2002	142.395,00	105.647,81	ABIERTA	0,047
18_A	10/04/2012	10/01/2003	284.850,87	143.059,09	ABIERTA	0,023
19_A	04/12/2009	18/12/2002	279.999,00	169.596,84	ABIERTA	0,105
20_A	07/05/2012	16/12/2002	104.353,44	84.653,80	ABIERTA	0,020
21_A	03/11/2011	13/01/2003	117.924,50	91.990,55	ABIERTA	0,038
22_A	06/05/2011	30/04/2003	204.500,00	164.245,54	ABIERTA	0,055
23_A	03/02/2009	21/04/2003	168.750,00	142.947,00	ABIERTA	0,134
24_A	05/03/2012	05/05/2003	105.000,00	80.480,64	ABIERTA	0,026
26_A	02/05/2012	30/04/2003	94.440,00	50.526,63	ABIERTA	0,020
27_A	19/06/2012	20/05/2003	151.110,00	120.273,09	ABIERTA	0,016
28_A	29/03/2011	27/06/2003	150.300,00	115.628,04	ABIERTA	0,059
29_A	07/05/2012	14/10/2003	150.256,00	108.233,94	ABIERTA	0,020
30_A	14/11/2012	25/09/2003	196.344,00	123.797,27	ABIERTA	0,002
31_A	03/11/2009	18/09/2003	150.813,45	78.306,48	ABIERTA	0,108
32_A	30/10/2012	30/10/2003	68.034,00	54.597,85	ABIERTA	0,003
33_A	05/09/2011	21/10/2003	89.400,00	72.178,90	ABIERTA	0,043

As we have movements in different dates that we have to compare, we'll have to use an IR to discount them. For this task we have this table:

CURVA DE DESCUENTO CON EL TIPO DE INTERÉS A APLICAR SEGÚN NÚMERO DE DÍAS

F_DATOS	FECHA	N_DIAS	TIPO
30/11/2012	04/12/2012	1	0,0034
30/11/2012	04/01/2013	31	0,0034
30/11/2012	04/02/2013	62	0,0034
30/11/2012	04/03/2013	90	0,0034
30/11/2012	04/04/2013	121	0,0034
30/11/2012	06/05/2013	153	0,0035
30/11/2012	04/06/2013	182	0,0035
30/11/2012	04/07/2013	212	0,0035
30/11/2012	05/08/2013	244	0,0035
30/11/2012	04/09/2013	274	0,0035
30/11/2012	04/10/2013	304	0,0035
30/11/2012	04/11/2013	335	0,0035
30/11/2012	04/12/2013	365	0,0035
30/11/2012	06/01/2014	398	0,0035
30/11/2012	04/02/2014	427	0,0036
30/11/2012	04/03/2014	455	0,0036
30/11/2012	04/04/2014	486	0,0036
30/11/2012	05/05/2014	517	0,0037
30/11/2012	04/06/2014	547	0,0037
30/11/2012	04/07/2014	577	0,0038
30/11/2012	04/08/2014	608	0,0039
30/11/2012	04/09/2014	639	0,0039
30/11/2012	06/10/2014	671	0,0040
30/11/2012	04/11/2014	700	0,0041
30/11/2012	04/12/2014	730	0,0042
30/11/2012	04/12/2015	1095	0,0052
30/11/2012	05/12/2016	1462	0,0068
30/11/2012	04/12/2017	1826	0,0086

The last table contains the data on the devaluation of the house prices:

Series de precios en España para vivienda y suelo

Trimestre	Precio (€)		
	Vivienda Libre	Suelo Urbano	Suelo P.Juridicas
2000 1T	856,20		
2000 2T	879,80		
2000 3T	891,60		
2000 4T	893,30		
2001 1T	930,30		
2001 2T	962,40		
2001 3T	982,60		
2001 4T	992,70		
2002 1T	1.051,70		
2002 2T	1.117,40		
2002 3T	1.142,70		
2002 4T	1.164,60		
2003 1T	1.230,30		
2003 2T	1.309,60		
2003 3T	1.344,90		
2003 4T	1.380,30		
2004 1T	1.456,20	206,47	243,03
2004 2T	1.538,80	226,40	220,60
2004 3T	1.570,80	227,07	223,13
2004 4T	1.618,00	247,31	252,29
2005 1T	1.685,40	258,57	283,61
2005 2T	1.752,80	254,51	212,19
2005 3T	1.781,50	263,85	268,62
2005 4T	1.824,30	267,25	254,12
2006 1T	1.887,60	257,41	258,19
2006 2T	1.942,30	258,42	166,55
2006 3T	1.956,70	273,65	246,97
2006 4T	1.990,50	284,60	253,72
2007 1T	2.024,20	271,78	292,52
2007 2T	2.054,50	280,60	316,86
2007 3T	2.061,20	285,00	324,60
2007 4T	2.085,50	277,03	321,34
2008 1T	2.101,40	250,90	328,66
2008 2T	2.095,70	258,82	305,27
2008 3T	2.068,70	257,07	342,88

§1 Resolution of the problem

§1.1 Calculate the probabilities of foreclosure and failure

We calculate the probability that a case ends in foreclosure/failure and is open after the T_i^{th} quarter by using the definition of conditional probability,

$$\mathbb{P}(\text{Case ends in } X | \text{Open at } T_i) = \frac{\mathbb{P}(\text{Case ends in } X \cap \text{Open at } T_i)}{\mathbb{P}(\text{Open at } T_i)}.$$

§1.2 Estimate LGD using LTV

For an open case, we do not know what the LGD will be, and we must estimate it using other parameters (which are known).

In this section we will show that for cases ending in foreclosure there is a very strong correlation between the LGD and LTV (Correlation coefficient =0.99850).

The LTV is easy to calculate, so we can assume it will be a good estimate for LGD.

LGD

We calculate LGD using the following formula,

$$LGD_i = 1 - \frac{\sum R_i - \sum G_i}{EAD_i} + G_{indirectos}$$

Where

- 1) The R_i are the payments made to the bank by the borrower (For cases ending in foreclosure this includes the repossession of the property).
- 2) The G_i are costs incurred by the bank in the recovery process (repossession and litigation costs etc.).
- 3) The EAD is the Exposure at Default, and is total debt owed by the borrower to the bank at the default date.
- 4) $G_{indirectos}$ are indirect costs incurred by the bank.

The value of each transaction R_i and G_j can be calculated in real terms at the default date using the discounting formula. This assumes that the interest rate (Risk Free Rate) is constant.

$$\text{Flujo}(\text{descontado}) = \frac{\text{importe del flujo}}{(1 + \text{tasa})^{\frac{(F(\text{entrada}) - F(\text{flujo}))}{365}}}$$

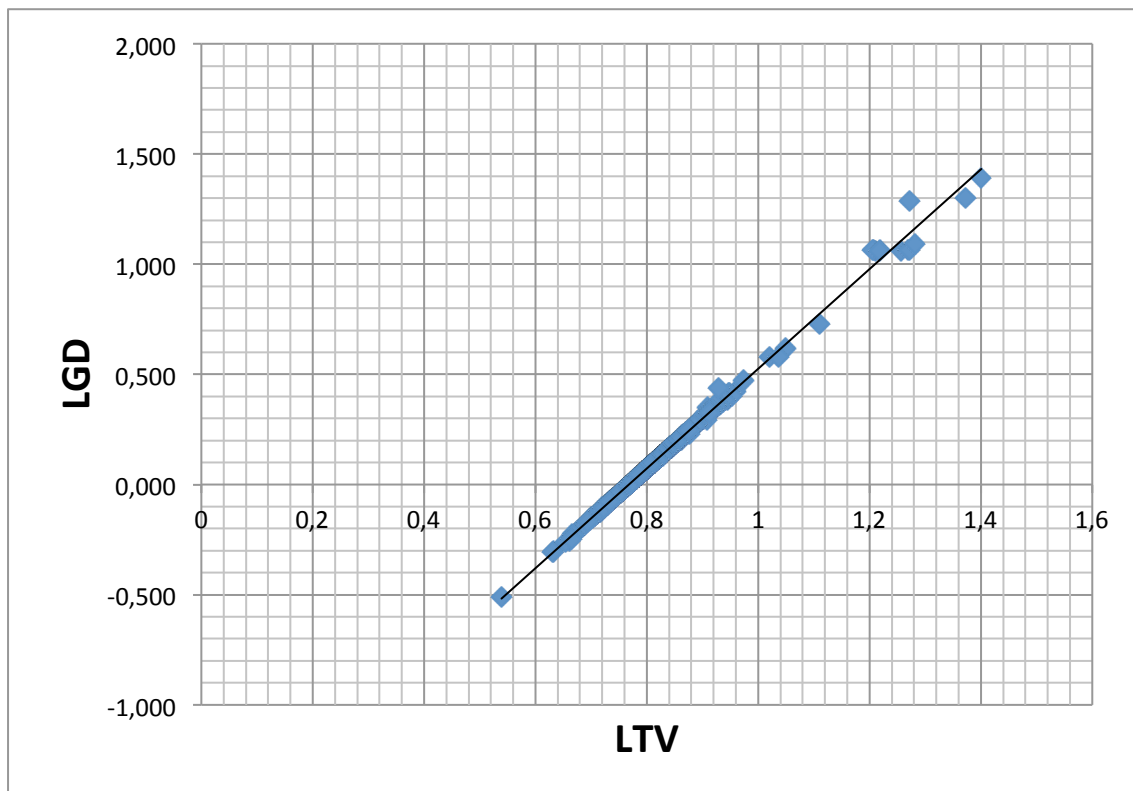
LTV

The LTV (Loan to Value ratio) is the ratio of the EAD to the value of the original value of the property.

$$LTV = \frac{EAD}{\text{Initial Value of Property}}$$

Relationship between LGD and LTV

We plot the LGD against the LTV of each case ending in foreclosure.



LGD and LTV are highly correlated (Correlation coefficient = 0.9850), and the regression line is given by

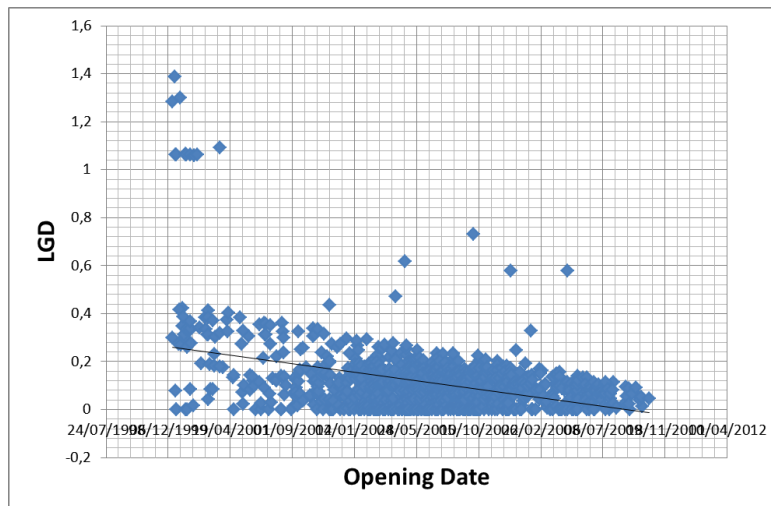
$$LGD = 2,2628 LTV - 1,7374$$

So, we take this as our estimate for the LGD for cases which have not yet closed.

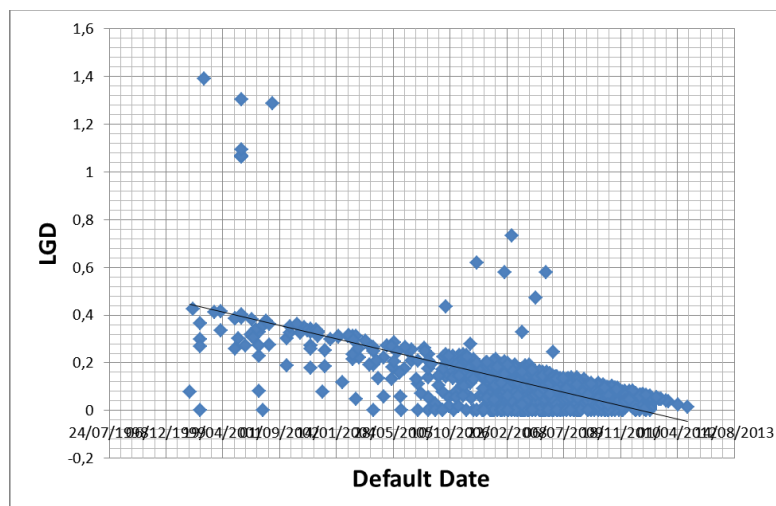
Relationship between LGD and other variables

We have also taken into consideration the fact that the LGD could have a high correlation with variables other than the LTV. In order to check this possibility, we have proposed some regression models which will show whether this correlation is high or not.

First of all, we present the relationship between the opening date and the LGD. This is unreliable as we expect cases which are quickly resolved to have a lower LGD, which may explain the apparent decrease in LGD with opening date.



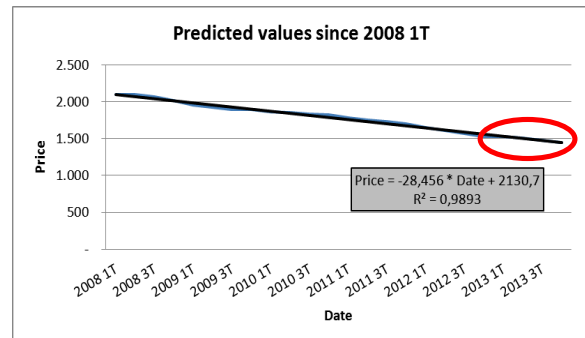
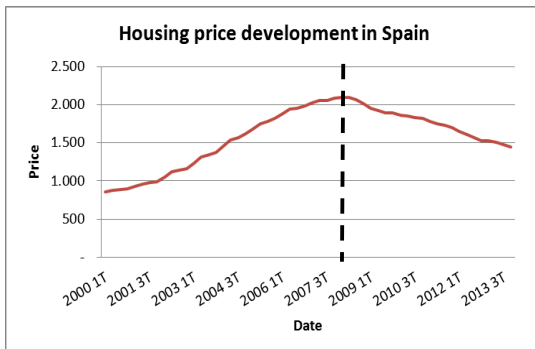
Secondly, we have also studied how the default date affects the LGD. We have found similar results, as we may see in the next picture.



We have also concluded that the default date is not a reliable estimator for the LGD for the same reason that the opening date.

We also ask: how does the current housing price trend affect LGD? It doesn't

Date	Price prediction
2013 1T	1.533,10 €
2013 2T	1.504,65 €
2013 3T	1.476,19 €
2013 4T	1.447,73 €



The housing price evolution, as well as its different trends in the last years, exemplify the life cycle of every housing bubble. We wonder whether the expected current value of the house in the LTV formula is useful so that we can predict the LGD. Unfortunately, the correlation was weaker.

LGD estimation

As we saw above, we are going to use the LGD formula in order to calculate it for closed contracts.

Let us remember that the main goal of these pages is obtaining a method which will allow us to estimate the LGD of those contracts that are still open. How are we doing this?

We will calculate the expected LGD for each case k that have been in default for T quarters (trimestres). We will use the law of total expectation, and partition the cases into those ending in foreclosure, failure and being cured.

$$\begin{aligned}
 \mathbb{E}(LGD_k | \text{Open at } T_i) &= \mathbb{P}(\text{Foreclosure} | \text{Open at } T_i) \mathbb{E}(LGD_k | \text{Foreclosure}) \\
 &+ \mathbb{P}(\text{Cured} | \text{Open at } T_i) \mathbb{E}(LGD_k | \text{Cured}) \\
 &+ \mathbb{P}(\text{Failed} | \text{Open at } T_i) \mathbb{E}(LGD_k | \text{Failed})
 \end{aligned}$$

The expectations of LGD_k on the right hand side are assumed to be independent of whether the case is open at T_i . We also assume that

- 1) $\mathbb{E}(LGD_k | \text{Cured}) \approx 0$, as the bank will not incur any recovery/litigation costs.
- 2) $\mathbb{E}(LGD_k | \text{Failed}) = 1$, as the bank must absorb the entire debt.

Therefore,

$$\mathbb{E}(LGD_k | \text{Open at } T_i) \approx \mathbb{P}(\text{Foreclosure} | \text{Open at } T_i) \mathbb{E}(LGD_k | \text{Foreclosure}) + \mathbb{P}(\text{Failed} | \text{Open at } T_i) \text{ for } i, k = 1, 2, \dots$$

So, to calculate the Expected LGD for each k we need to do the following

- 1) Calculate the probabilities of ending in foreclosure and failure given that the case is open at the end of the T_i^{th} quarter.
- 2) Find an estimate for $\mathbb{E}(LGD_k | \text{Foreclosure})$. To do this we will use the very high correlation between LGD and LTV for cases ending in foreclosure.

One could ask: how are those probabilities estimated? We will use Bayes theorem and Conditional Probability definition. So, for example, to calculate the probability of ended up in foreclosure given that the contract was open after T_i we use:

$$\begin{aligned} P(\text{Foreclosure} | \text{open } T_i) &= \frac{P(\text{Foreclosure} \cap \text{open } T_i)}{P(\text{open } T_i)} \\ &= \frac{P(\text{open } T_i | \text{Foreclosure}) \cdot P(\text{Foreclosure})}{P(\text{open } T_i)} \end{aligned}$$

In particular:

$$P(\text{Foreclosure}) = \frac{\# \text{ Closed contracts that ended up in Foreclosure}}{\# \text{ Closed contracts}}$$

$$P(\text{open } T_i) = \frac{\# \text{ Closed contracts that were open after } T_i \text{ quarters}}{\# \text{ Closed contracts}}$$

$$\begin{aligned} P(\text{open } T_i | \text{Foreclosure}) &= \frac{\# \text{ Closed contracts as foreclosure that were open after } T_i \text{ quarters}}{\# \text{ Closed contracts as foreclosure}} \end{aligned}$$

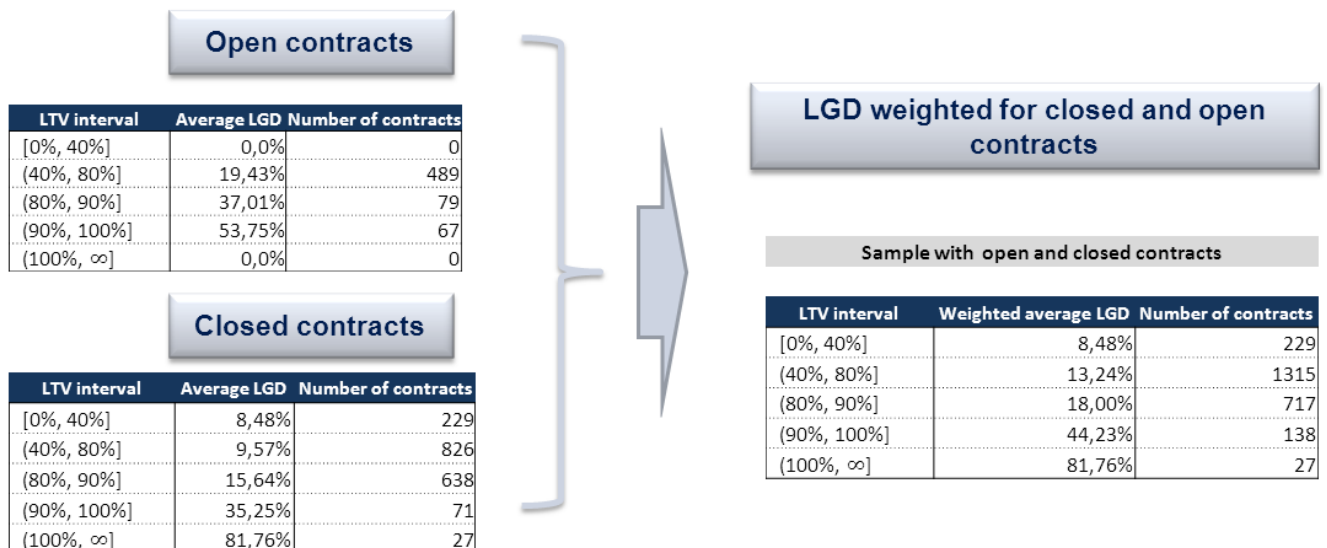
§2. Applications of LGD

Calculate LGD with buckets of LTV

We have divided the possible values of LTV in five intervals or buckets to present some results for each one, in particular the average LGD.

For this task, we have considered two different scenarios.

On the one hand, we consider the original valuation as the value of each house so we obtain:



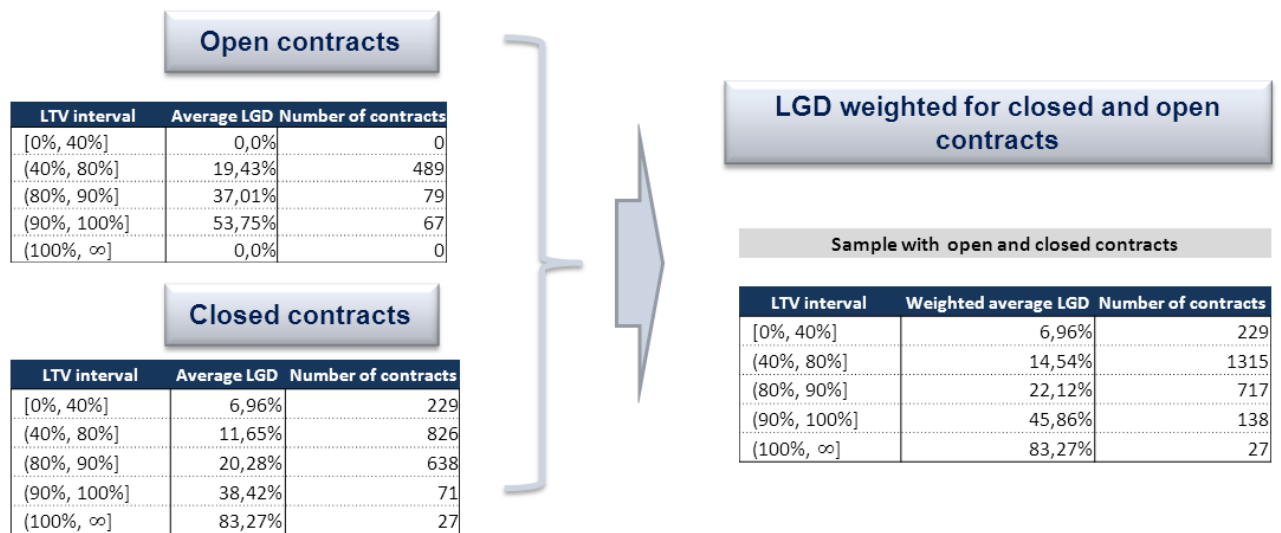
We see the buckets for both open and closed contracts. For each interval of LTV we have the average of LGD and the number of contracts of our database. Also, we have a table with the open and closed contracts together.

We notice that the bucket in which more contracts are included is the second one, where the LTV is between 40% and 80%.

In addition, we can see that the LGD has an increasing trend. This is reasonable because in buckets we group the contracts and the bigger LTV is, the bigger the LGD will be, as we saw before in the regression line.

Finally, we checked that there is no contract for which LGD=100%. This is because all of our open contracts are young so there is no failed contract yet.

On the other hand, we consider a stressed scenario in which we have discounted the movements with a 4% ratio which corresponds to an average interest rate before the crisis.



In this scenario, we can see that the number of contracts is the same than considering the mean ratio. This is because the LTV doesn't depend on the discounting rate. LTV only depends on EAD divided by valuation.

Moreover the average LGD of each bucket is higher than in the previous scenario because the rate interest is bigger.

Finally, we consider a scenario in which the interest rate was dependent on the time but the results obtained were similar to the ones obtained with the average interest rate.

Calculate Expected and Unexpected Losses

The LGD parameter is necessary to estimate the losses that a bank must affront.

The expected losses, represented as EL, are calculated via:

$$EL = PD \cdot LGD \cdot EAD$$

where PD represents the probability of counterparty's default.

These losses are covered by provisions which are resources that the entity must keep until the payment is satisfied.

Unexpected losses are calculated using a formula that appears in *DV1 Estimaciones de las LGD downturn en las carteras hipotecarias en España*, a paper from the Bank of Spain. This formula

estimates unexpected losses using the following variables: PD, LGD, EAD and M (Maturity of the contract).

Calculating unexpected losses is very important for entities because they must have Regulatory Capital to cover the unexpected losses.

§3. Conclusions

Looking back at all the work we have done, we can conclude:

- The LGD depends strongly on LTV: the correlation coefficient between the LTV and the LGD (considering the foreclosure ending contracts) is 0.9850 which means that LTV is a good predictor for LGD.
- Other variables have been considered to predict LGD: the opening date, the number of days between opening date and default date and the updated LTV. However these are not good variables for forecasting LGD.
- For both the higher interest rate and the risk free curve, LTV has still a high correlation with LGD.
- In the stressed scenario, we predict an increase in LGD.

§4. Next steps

There are some extensions that could be taken into account to improve our project. For example:

- Study correlation between LGD and other variables (different time periods, maturity of the contract...).
- Include several variables in the model to improve the forecast of LGD for open contracts. For example multiple regression models, time series models ...
- Compare results with those for other interest rates (e.g. highest, lowest, regressions).
- Calculating the LGD parameter for others different segmentation buckets of LTV
- Calculate the joint LGD to operations of the points above with different assumptions
- Calculate the associated capital requirements for Spanish financial institutions considering the obtained results.