Household Debt and Economic Growth in Europe

L. Barbaglia,[†] S. Manzan,[*‡] and E. Tosetti[¶]

†Joint Research Centre, European Commission, Ispra, Italy
‡Bert W. Wasserman Department of Economics and Finance, Baruch College, New York, USA
¶University of Padua, Padua, Italy
*Corresponding author. Email: sebastiano.manzan@baruch.cuny.edu

ONLINE APPENDIX

Appendix A: European Datawarehouse preparation

We cleansed the data set by eliminating records with missing entries, errors as well as duplicated entries. We only kept data on borrower's gross income if the income has been verified by the bank, rather than self reported, to avoid fraudulently overstated income, a problem that has been pointed out by Mian and Sufi (2017) for the US case. We also dropped records in the first and last percentiles of the continuous variables in our analysis, calculated for each year and each country in the sample. In addition, we consider only loans originated after the year 2000 for the purchase of a property, with interest rate type from 1 to 4 (fixed for life, fixed with resets, floating for life, and floating with resets), that have a Loan-to-Value (LTV) ratio smaller than 130, household income larger than 10 thousand Euro annually, and interest rate between 0 and 15%. The outcome of this filtering is a data set of over 4.6 million loans expressed in Euro across 8 European countries, namely, Belgium, France, Ireland, Italy, the Netherlands, Portugal, Spain and the UK[14]. We aggregate these data at the level of the region where the asset underlying each loan is located using the NUTS3 regional classification. To obtain reliable NUTS3 level aggregates, we drop all observations in regions for which the number of loans observed at each point in time is below the threshold of 50 loans.

The interest rate provided in the data set refers to the rate currently charged on the loan, rather than the rate applied at origination. In our application we are interested in the interest rate at origination and we use the current rate to reconstruct it using the following rules. If the loan has a rate defined as "floating for life", we consider the current interest rate as the rate at origination if it is within 1% of the average mortgage rate (for maturities less than 5 years) published by the ECB for the country. Otherwise we multiply the current interest rate by the ratio of the current ECB average mortgage rates and its value at the origination date. In case the floating rate is linked to an index we calculate the current spread and add it to the value of the index at origination to obtain the interest rate. For fixed rate loans we compare the current rate on the loan to the ECB average rate on maturities of 5 years or longer at origination. We set the rate at origination equal to the current rate if it is within 1% of the ECB rate for the country otherwise we rescale it using the benchmark rates. Although the "fixed for life" category suggests that the current interest rate should also be the rate at origination, there is large degree of inconsistency between the values reported in the data set and the prevalent rates at the time the mortgage was originated.

Appendix B: Sample representativeness

This Section addresses the relevant question of how representative the European Datawarehouse (ED) is for the financial situation of European households (Gaudêncio, Mazany, and Schwarz 2019). Table 5 reports the summary statistics for the sample size at the level of NUTS3 region by country. The median number of observations per region ranges between 213 and 2,568, and the minimum of the first quartile equal to

14. The participation of German banks to this ECB operation has been limited and it is possibly associated to low German home-ownership rate (Voigtländer 2009). The lack of a sufficient number of loans for many NUTS regions has lead to exclude Germany from the analysis.

120 while the maximum of the third quartile is 4,857. Ideally, one would examine the representativeness of the data set by computing the percentage of total loans that is covered in the ED data set. We could find reliable figures on the number of new loans granted each year by country only for Spain[15]. In this case, we observe that the ED data set covers approximately 15% of the total number of mortgages granted over the period 2003-2013. We next investigate the issue of how representative our sample is for the underlying population by comparing the key variables included in our analysis to the same variables constructed from consumer finance household surveys

There are several concerns that the loans submitted to the ECB for these financing operations might not be representative of the underlying population. First, the ECB sets a threshold on the credit quality of the ABS which requires banks to include in the pool high quality loans. In this sense the sample might provide a sample that is of significantly higher credit quality relative to the overall level in a country and under-represent low credit quality loans (that are more likely to have higher DTI and default rates). Second, banks participating more actively to the ECB liquidity operations might be those with a precarious financial situation and unable to access capital markets. In this sense the pool of loans that they submit might be larger and of lower quality relative to the population. Both of these effects can potentially bias our results, although in opposite directions. To evaluate to what extent our sample is representative of the underlying household financial situation, we consider the Household Finance and Consumption (HFCS) survey[16] that is coordinated by the ECB and provides a standardized set of variables across European countries (except for the UK that is not part of the Euro-area). With respect to the interest rate, we obtained the average interest rate on new residential loans from the European Mortgage Federation (EMF)[17]. Unfortunately, both the HFCS and EMF provide information only at the country-level rather than for regions. Hence, we compare the key variables obtained from ED and the HFCS survey at the national level, with the caveat that the representativeness at the regional level remains to be demonstrated due to lack of information in the Survey.

We focus the evaluation of the representativeness of the ED data set on three important variables that are used in our empirical analysis, namely, the DTI, house prices and the loan interest rate. For the HFCS, we construct a national weighted average of DTI and house prices over time, where observations are weighted to ensure the representativeness of the Survey[18]. Figure 9 shows the number of observations in ED data versus the HFCS survey by origination year of the loan. The ED data set has (relatively) fewer observations in the early 2000s while loans originated after 2010 are scarcer in the HFCS. For both ED and HFCS, the year with most observations is 2006, when we have approximately 500 thousand loans for ED and 1.1 thousand for the HFCS survey. In general, ED provides for each country and year thousands of loans as opposed to a few hundreds from the HFCS. In terms of the geographical

15. Total number of mortgages from 2003 published by the Spanish Statistical Office at https://www.ine.es/dyngs/INEbase/en/operacion.htm?c=Estadistica_C&cid=1254736170236&menu=ultiDatos&idp=1254735576757.
16. More information at https://www.ecb.europa.eu/stats/ecb_surveys/hfcs/html/index.en.html.
17. More information at https://hypo.org/.
18. The design weight is adjusted for non-response and ensures the representativeness of the survey.

distribution of the loans, the countries more represented in ED are the Netherlands, France and Spain while in HFCS they are France, Ireland, and Portugal. Figure 10 compares the temporal evolution of the median DTI in the two data sets, with the bands representing the first and third quartile of the cross-sectional distribution of the variable in that year. For most countries the median DTI and its distribution are quite similar both in level and evolution over time. The DTI calculated in the HFCS is more volatile due to the small sample of the Survey, while the ED vary more smoothly. The largest differences between the two data sets appears when considering Italy and the Netherlands. For Italy, the median DTI from ED is approximately 1.5 point higher on average relative to HFCS and its distribution seems to be shifted upwards. This might indicate that the pool of loans originated in Italy represents riskier loans than those included in the Survey. For the Netherlands, we find the opposite result that the median DTI of the ED loans is on average close to 2, while almost 3.5 for the loans in the HFCS survey.

In Figure 11 we show the median house price obtained from the HFCS survey and from ED. Similarly to the DTI ratio, the time series of house prices overlap for most countries, including Italy. Only for the Netherlands we find that the median house price is significantly higher for HFCS than ED (205 thousand euro vs 131 thousand on average over the sample period). Hence, loans originated in the Netherlands in ED have lower DTI ratio and higher house purchase price relative to those collected in the HFCS survey. This points to the fact that Dutch loans in ED might be of higher credit quality relative to the population since more expensive properties require bigger loans and larger incomes to keep the DTI ratio low. This feature might bias our results against finding a role for household leverage, given that the pool of Dutch loans are very high quality and Figure 10 shows no significant increase in leverage before the Great Recession. Finally, Figure 12 shows the average interest rate in ED and the average provided by the EMF. Overall, there is consistency between the rates both in level as well as in dynamic behavior.

Summarizing, graphs from Figures 9 to 12 seem to suggest the broad consistency of the main household finance variables in ED with alternative sources such as the HFCS. However, ED provides two significant advantages over using survey data. Firstly, it provides regional information that allows to analyze and model the within-country variability of the variables of interest. Secondly, the availability of several thousand loans per region-year delivers more robustness to the analysis.

Appendix C: Housing supply elasticity

Saiz (2010) proposes a measure of housing supply elasticity (HSE) that depends on the geographic characteristics of a region, such as terrain elevation and water bodies. In areas with few geographic constraints the housing supply is elastic and adjusts quickly to changes in demand with a small impact on house prices. Instead, prices will react significantly when geographical constraints are binding and housing supply cannot react fully to accommodate demand. An expansion of the mortgage supply has the effect of increasing housing demand and produces differential effects on house prices in elastic and inelastic regions. In particular, prices will increase rapidly in those regions with inelastic housing supply, thus requiring borrowers to apply for larger

Table 5. Descriptive statistics for the number of loans per year-NUTS3 region by country.

Country	25th perc.	Median	Mean	75th perc.
BE	431	926	1,639	2,185
ES	268	682	1,902	1,798
FR	346	773	1,228	1,610
IE	473	1,559	2,774	4,790
IT	136	302	509	580
NL	1,281	2,568	3,704	4,857
PT	150	354	840	858
UK	120	213	315	388

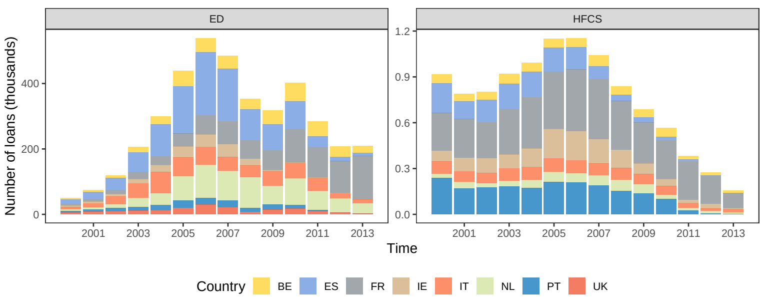

Figure 9. Number of loans (thousands) in the ED data set (left panel) and the HFCS survey (right panel) by country and year of origination.

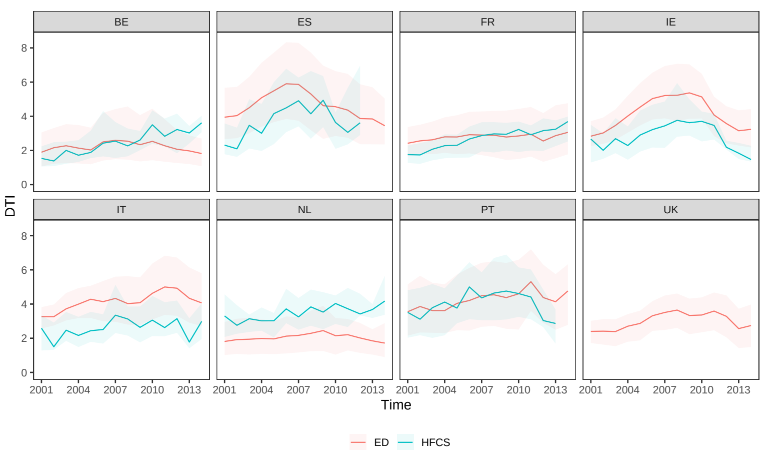

Figure 10. Median DTI calculated on all loans originated in a country-year using data from the ED and the HFCS. The bands represent the first and third quartiles of the cross-sectional distribution of DTI. The United Kingdom does not participate to the HFCS survey.

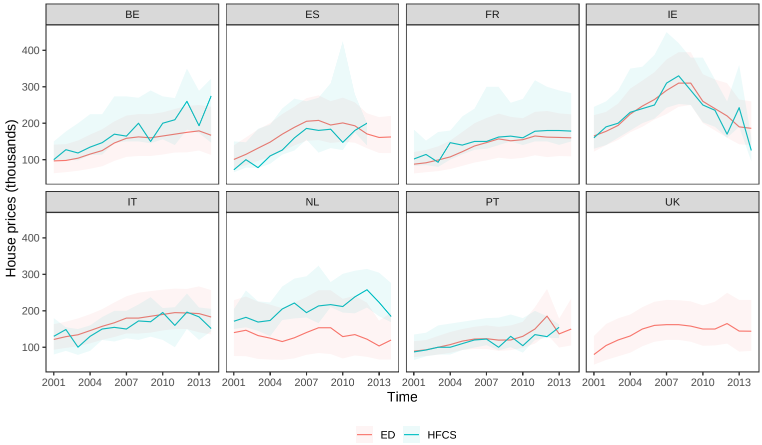

Figure 11. Median house price from all loans originated in a country-year using the ED and the HFCS data. The bands represent the first and third quartiles of the cross-sectional distribution of the house price. The United Kingdom does not participate to the HFCS survey.

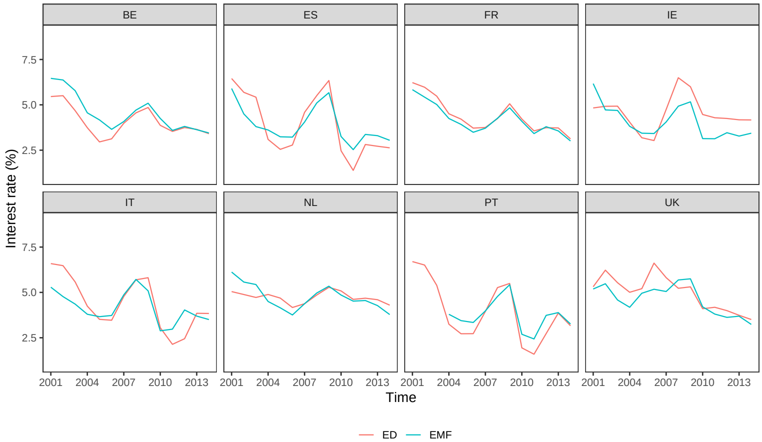

Figure 12. Average interest rate at loan origination in the ED data set (in red) and the averages reported by the European Mortgage Federation (in green) for a country-year.

loans: inelastic regions are characterized by higher house prices growth and household leverage relative to elastic regions. Mian and Sufi (2010) find empirical evidence that house prices increased significantly in US counties with relatively inelastic housing supply. Moreover, they show that these regions were the most severely hit by the slump in output and employment during the Great Recession.

To construct his measure of HSE, Saiz (2010) considers the following housing supply equation:

$$\Delta P_g = \alpha + \beta_g \Delta H_g + \sigma_g \Delta CC_g + \boldsymbol{R}_g + \epsilon_g, \tag{4}$$

where ΔP_g is the log-difference in housing prices in area g over the period from 1970 to 2000, H_g is the growth in the number of housing units, CC_g is the percentage growth in construction costs in the same period, and $\boldsymbol{R}_g$ is a set of regional dummies. In the above equation, β_g is the so-called *inverse housing supply elasticity*. It represents the price sensitivity to demand shocks, and is assumed to be a decreasing function of land availability. Accordingly, the author suggests to approximate β_g by:

$$\beta_g = \beta_1 + \beta_2 \mathrm{UL}_g, \tag{5}$$

where UL_g represents the share of land that is unavailable for residential development. Following Mian, Rao, and Sufi (2013), regions with larger β_g (thus with more unavailable land for building) have a relatively inelastic housing supply, for which we expect larger values of DTI. Accordingly, the authors use Equation (5) as instrument for the change in DTI in the IV estimation of Equation (1). We observe that β_g is a linear transformation UL_g and its variability across regions is only due to variation in UL_g. Hence, IV estimation results are identical whether one uses β_g or the share of unavailable land, UL_g. Given that the ED data set has few mortgages originated before 1996, we do not have reliable data on house prices over a long period of time: we overcome this shortcoming by instrumenting DTI in Equation (1) with UL_g, rather than β_g, and refer to such instrument as HSE. We expect land-constrained regions (i.e., with large UL_g values) to have a relatively inelastic housing supply and thus larger DTI.

We proxy the share of unavailable land using data from the JRC LUISA Territorial Modelling Platform[19]. Such data combine information from several sources, including satellite images on human and industrial settlements, data on elevation of the earth's surface, and information on protected areas where building is not permitted by the law. Data are available at each decennial census, for 1990, 2000 and 2010. We calculate the total area that is not available for building purposes by considering (i) areas that have already been built, (ii) non-buildable slopes, (iii) protected or green urban areas and (iv) water bodies. We obtain UL_g by dividing the total area that is not available for building calculated for the year 2000 by the total area in the NUTS3 region[20].

19. More information can be found at https://ec.europa.eu/jrc/en/luisa.

20. We observe that Saiz (2010) only restricts his calculation to the area within the 50km radii from the centroid of each metropolitan statistical areas, in order to capture the portion of land around cities that is not available for residential or commercial development. Although in our application we do not have information to identify the area surrounding each city, we observe that the NUTS3 classification allow a subdivision in relatively small-sized regions, covering one or more urban centers.

Figure 13 displays the map of unavailable land, showing substantial across-country heterogeneity. Regions with the largest constraints belong to the mountainous areas of the north of Italy and Spain, south of France as well as some densely populated large urban areas. Peripheral countries like Spain, Italy and Portugal are on average more land-constrained than the core countries UK, France, the Netherlands and Belgium, probably due to the physical constraints in the form of mountainous terrain for the first group of countries. The high percentage of constrained land that can be observed for the Netherlands may rather reflect the adoption of restrictive land use regulation (Vermeulen and Rouwendal 2007).

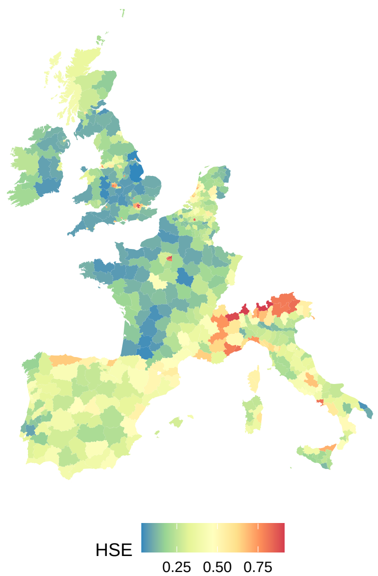

Figure 13. Map of the Housing Supply Elasticity (HSE) at the NUTS3 level.

Figure 14 shows the average growth of DTI and house prices relative to their levels in 2000, together with the median GDP in euro. In these graphs we split the regions in those with an inelastic housing supply (top quartile of the HSE distribution), versus those with an elastic supply (bottom HSE quartile). Leverage for these two groups grows at a similar pace between 2003 and 2007, while it diverges after 2008 when the DTI for the inelastic regions continues to grow until 2011. Instead, the house prices in both elastic and inelastic regions grew at a similar rate to about 40% until 2007, followed by a decline of approximately 20% for the inelastic regions and relatively stable prices in the elastic regions. As for GDP, it is interesting to observe that inelastic regions are characterised by a median GDP consistently higher than elastic regions. Overall, the HSE does not seem to be able to correctly identify those regions that increase leverage, although the increase is not driven by the diverging dynamics of the housing market in the two groups. Under the approach advanced by Mian and Sufi for the US, we should expect that areas with relatively small percentage of land available for building are those that registered the highest growth in house prices, while this does not seem to be the case for Europe. One possible explanation for this result is that, at least for Europe, the regions with high land constraint are mostly scarcely populated regions located in the mountainous regions of north Italy, Spain and southern France.

Overall, the HSE does not seem to be an appropriate instrument for DTI for the European case for a number of reasons. First, in Europe areas with relatively small percentage of land available for building are a mix of wealthy, high densely populated region with tough laws for home builders, such as the Netherlands, and poor, sparsely populated areas with severe land constrains due to mountainous territories, such as the north of Spain or the South of Italy. We also observe that the large differences in population density across European countries may mask important intra-country differences, thus making the HSE unreliable. For instance, Belgium and the Netherlands have almost four time the population density of Spain or Portugal. Finally, the validity of such an instrument has been questioned also for the US case: among others, Davidoff (2016) claims that strictly regulated regions are not only expensive because building in those areas is costly, but also because they offer better employment opportunities, thus attracting high-income workers. Given these motivations, we next introduce an alternative instrument for DTI that we believe better captures the credit supply shock in Europe.

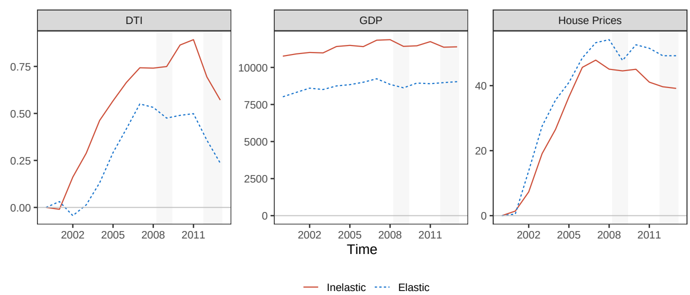

Figure 14. Change (relative to year 2000) of DTI and percentage growth of real GDP (in Euro) and house prices for the top and bottom quartiles of the distribution of the HSE instrument. The top quartile is denoted as *inelastic* and the bottom quartile as *elastic*.

Appendix D: Lending Standards

We consider some loan and borrower characteristics to evaluate their behavior between 2000 and 2010. In Figure 15 we show the loan balance at origination (in euro), the Debt-to-Income ratio, the house prices, the borrower's income (in euro), the loan's interest rate, the Loan-to-Value ratio, and the loan term (in months). These values are averages across all loans originated in a country in a certain year.

The overall picture that emerges from the Figure is that most countries experienced a large increase in average house prices and loan balances, in particular in Ireland and Spain. The trend was already in place in the early 2000s, but accelerated significantly after 2003 reaching a peak in 2006-2007. The similar growth in these two variables lead to a LTV ratio that was relatively stable in most countries. However, household incomes grew in these years but not fast enough to keep up with house prices and loan balances. This caused the DTI ratio to increase significantly, in particular in those countries with larger increases in house prices. The affordability of larger mortgages to purchase more expensive houses was maintained by a combination of lower interest rates and, in most countries, significantly longer loan terms.

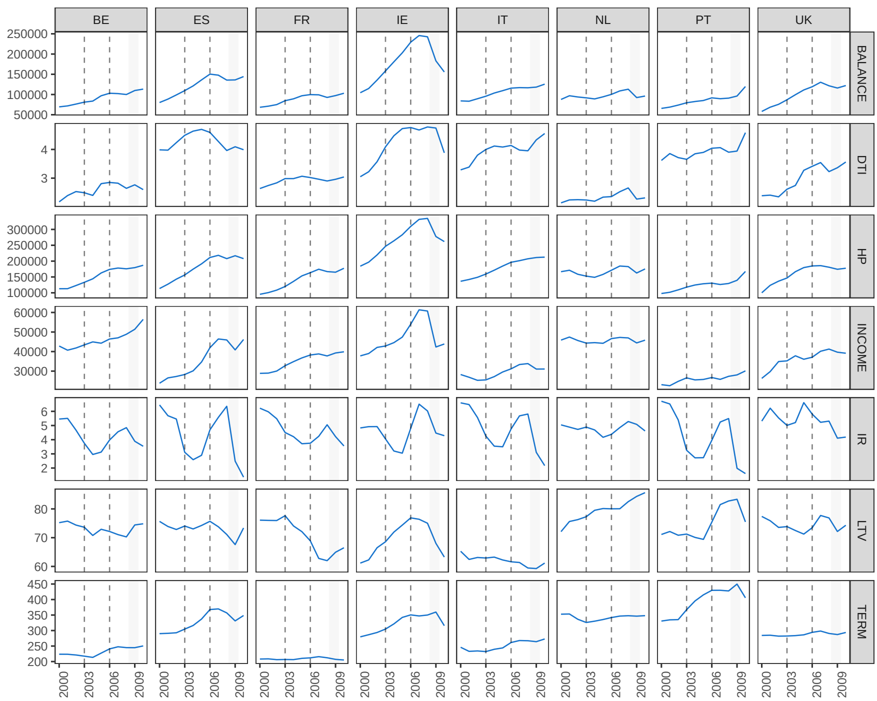

Figure 15. Variation over time of average loan and borrower characteristics: *BALANCE* represents the loan balance at origination (unit: euro), *DTI* the Debt-to-Income ratio, *HP* the house price, *INCOME* is the annual borrower income (unit: euro), *IR* the interest rate at origination, *LTV* the Loan-to-Value ratio, and *TERM* the duration of the loan (unit: months). The vertical bars are for 2003 and 2006, while the gray shaded area represents the CEPR recession period.

Appendix E: Estimation results of the credit shock model

Table 6 provides the estimation results for Equation (2) in the paper.

Table 6. Estimation results of Equation (2)

Variable	Country	Estimate	SE	t-stat	p-value
Intercept	BE	6.984	0.385	18.156	0.000
Intercept	ES	-11.852	0.422	-28.062	0.000
Intercept	FR	-5.078	0.452	-11.230	0.000
Intercept	IE	-8.328	0.495	-16.832	0.000
Intercept	IT	-2.098	0.452	-4.642	0.000
Intercept	NL	-5.188	0.442	-11.747	0.000
Intercept	PT	1.266	0.559	2.267	0.023
Intercept	UK	-9.249	0.447	-20.688	0.000
DTI	BE	-0.064	0.035	-1.842	0.066
DTI	ES	0.545	0.030	17.915	0.000
DTI	FR	0.142	0.033	4.373	0.000
DTI	IE	0.193	0.045	4.253	0.000
DTI	IT	-0.030	0.039	-0.775	0.438
DTI	NL	0.053	0.030	1.774	0.076
DTI	PT	-0.062	0.044	-1.408	0.159
DTI	UK	-0.228	0.037	-6.187	0.000
DTI^2	BE	0.002	0.004	0.620	0.535
DTI^2	ES	-0.041	0.004	-10.398	0.000
DTI^2	FR	-0.009	0.004	-1.925	0.054
DTI^2	IE	0.003	0.006	0.547	0.584
DTI^2	IT	0.003	0.005	0.554	0.580
DTI^2	NL	-0.005	0.004	-1.082	0.279
DTI^2	PT	0.009	0.005	1.593	0.111
DTI^2	UK	0.066	0.006	11.746	0.000
INCOME (log)	BE	-0.296	0.034	-8.767	0.000
INCOME (log)	ES	1.222	0.025	48.167	0.000
INCOME (log)	FR	0.400	0.029	13.804	0.000
INCOME (log)	IE	0.691	0.035	19.783	0.000
INCOME (log)	IT	0.182	0.029	6.258	0.000
INCOME (log)	NL	0.309	0.028	10.972	0.000
INCOME (log)	PT	0.121	0.045	2.694	0.007
INCOME (log)	UK	0.818	0.029	28.361	0.000
LTV	BE	0.001	0.002	0.376	0.707
LTV	ES	-0.010	0.002	-4.127	0.000
LTV	FR	-0.002	0.003	-0.888	0.375
LTV	IE	-0.006	0.003	-1.657	0.098
LTV	IT	0.025	0.002	10.559	0.000
LTV	NL	0.001	0.002	0.379	0.705
LTV	PT	0.010	0.004	2.768	0.006
LTV	UK	0.039	0.003	11.749	0.000
LTV^2	BE	0.000	0.000	0.213	0.831
LTV^2	ES	-0.000	0.000	-0.472	0.637
LTV^2	FR	0.000	0.000	1.120	0.263
LTV^2	IE	0.000	0.000	1.335	0.182
LTV^2	IT	-0.000	0.000	-10.645	0.000
LTV^2	NL	-0.000	0.000	-0.585	0.558
LTV^2	PT	-0.000	0.000	-1.526	0.127
LTV^2	UK	-0.000	0.000	-17.287	0.000
RATE 2	BE	0.357	0.544	0.657	0.511
RATE 2	ES	-1.859	0.544	-3.419	0.001

RATE 2	FR	-0.822	0.551	-1.492	0.136
RATE 2	IT	-1.695	0.544	-3.117	0.002
RATE 2	NL	0.582	0.546	1.065	0.287
RATE 2	PT	-3.268	0.544	-6.003	0.000
RATE 3	BE	0.095	0.304	0.313	0.754
RATE 3	ES	-0.708	0.315	-2.250	0.024
RATE 3	FR	0.284	0.315	0.900	0.368
RATE 3	IE	0.114	0.307	0.371	0.711
RATE 3	IT	0.022	0.305	0.073	0.942
RATE 3	NL	1.329	0.348	3.820	0.000
RATE 3	PT	-2.541	0.326	-7.787	0.000
RATE 4	BE	-1.186	0.305	-3.889	0.000
RATE 4	ES	-1.780	0.316	-5.629	0.000
RATE 4	FR	2.108	0.330	6.384	0.000
RATE 4	IE	2.169	0.311	6.985	0.000
RATE 4	IT	-0.248	0.307	-0.808	0.419
RATE 4	NL	2.212	0.307	7.194	0.000
RATE 4	PT	-1.533	0.583	-2.631	0.009
TERM	BE	0.001	0.000	2.732	0.006
TERM	ES	-0.003	0.000	-9.419	0.000
TERM	FR	0.002	0.000	7.091	0.000
TERM	IE	-0.000	0.000	-1.401	0.161
TERM	IT	0.001	0.000	4.240	0.000
TERM	NL	0.000	0.000	0.430	0.667
TERM	PT	-0.002	0.000	-7.019	0.000
TERM	UK	-0.000	0.000	-0.182	0.856
BALANCE (log)		-0.027	0.027	-1.003	0.316
LIBOR		0.510	0.004	113.423	0.000
Observations	176,573				
R^2	0.311				
Adjusted R^2	0.311				

Interest rate type: 1: floating, 2: tracker, 3: fixed for life, or 4: fixed with future periodic resets.