

Unpublished Appendix: Ancillary Results and Robustness Checks on a Probit Model of Irish Mortgage Defaults

Gregory Connor and Thomas Flavin
National University of Ireland, Maynooth
September 2014

This unpublished appendix provides ancillary empirical results and some simple robustness checks on the probit model of Irish mortgage defaults presented in Connor and Flavin (2014) hereafter referred to as CF. The document is organized in three sections; first we present our robustness tests; second, we include ancillary tables to complement the results reported in CF; and third, we discuss the absence of “cures” (that is, successful workouts) in our sample.

1. Robustness Tests

In this section we compare the CF probit-based coefficient estimates with coefficient estimates using logit and using a simple linear-probability ordinary least squares approach. We show estimation details in each case. We plot the residuals from the linear-probability ordinary least squares estimates to check for heteroskedasticity. The data is identical to that in CF, and we follow the variable name conventions from that paper.

Recall that CF uses three subsample categories of mortgages: all loans, home loans, and buy-to-let loans, and five explanatory variables in the main estimation model: application affordability ratio, current affordability ratio, application loan-to-value, current loan-to-value, and log income. We apply three estimation methods: probit, logit and linear-probability ordinary least squares (see Greene (2008, pp. 770-794)).

The linear probability model has a number of shortcomings. This model serves as a useful robustness check rather than as sensible data generating process; we refer the interested reader to Greene (2008, pp 772-773) and references therein. For this model we calculate Eicker-White heteroskedasticity-consistent standard errors and the coefficient t-statistics are based upon these heteroskedasticity-consistent standard errors.

All the models were estimated using RATS and in all cases the nonlinear search routine (for probit and logit) converged quickly, as shown in the results below. See Enders (2003) for technical details on RATS estimation algorithms. The estimation results are shown in the following nine panels (for the three subsamples and three estimation methods). The results are quite uniform across the three subsamples and three estimation techniques. Each of the coefficients has the same sign in every single case of the nine cases: AppNet negative, Net positive, AppLTV negative, LTV positive, LogIncome negative. The probit and logit coefficients estimates are similar in magnitude within each subsample. The linear-probability ordinary least squares estimates are not comparable to probit/logit due to a nonlinear transformation between the models, but they do preserve the signs. Of the fifty-four coefficient estimates (six coefficients including constant time three estimation methods times three subsamples) only three have t-statistics less than 1.96 in magnitude – and these three are all for the same variable and subsample: the AppNet variable in the buy-to-let subsample is insignificant using all three estimation methods.

The maximum number of iterations needed for convergence is five, which is low, and this indicates that the empirical likelihood surface is well-behaved.

Probit Estimation for All Loans

Binary Probit - Estimation by Newton-Raphson

Convergence in 4 Iterations. Final criterion was 0.0000000 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations 24993
 Degrees of Freedom 24987
 Skipped/Missing (from 28377) 3384
 Log Likelihood -16375.1310
 Average Likelihood 0.5193441
 Pseudo-R² 0.0729950
 Log Likelihood(Base) -17296.8518
 LR Test of Coefficients(5) 1843.4417
 Significance Level of LR 0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	1.195719666	0.153189147	7.80551	0.0000000
2. APPNETS	-0.178311566	0.070534686	-2.52800	0.01147149
3. NETS	0.338455435	0.034542978	9.79810	0.0000000
4. APPLTV	-0.946680026	0.047507693	-19.92688	0.0000000
5. LTV	0.875840399	0.026224509	33.39778	0.0000000
6. LOGINCOME	-0.217966364	0.018730206	-11.63716	0.0000000

Logit Estimation for All Loans

Binary Logit - Estimation by Newton-Raphson

Convergence in 4 Iterations. Final criterion was 0.0000011 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations 24993
Degrees of Freedom 24987
Skipped/Missing (from 28377) 3384
Log Likelihood -16370.3212
Average Likelihood 0.5194440
Pseudo-R² 0.0733718
Log Likelihood(Base) -17296.8518
LR Test of Coefficients(5) 1853.0612
Significance Level of LR 0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	2.028625662	0.251509324	8.06581	0.00000000
2. APPNETS	-0.297547130	0.116727572	-2.54907	0.01080096
3. NETS	0.550564290	0.057019151	9.65578	0.00000000
4. APPLTV	-1.571361416	0.079445454	-19.77912	0.00000000
5. LTV	1.435511074	0.044044920	32.59198	0.00000000
6. LOGINCOME	-0.363598531	0.030830923	-11.79331	0.00000000

Linear-Probability Ordinary Least Squares Estimation for All Loans

Linear Regression - Estimation by Least Squares

With Heteroscedasticity-Consistent (Eicker-White) Standard Errors

Dependent Variable DEFAULT

Usable Observations	24993
Degrees of Freedom	24987
Skipped/Missing (from 28377)	3384
Centered R ²	0.0707082
R-Bar ²	0.0705223
Uncentered R ²	0.5137699
Mean of Dependent Variable	0.4767734966
Std Error of Dependent Variable	0.4994702305
Standard Error of Estimate	0.4815363792
Sum of Squared Residuals	5793.9177071
Log Likelihood	-17196.3084
Durbin-Watson Statistic	1.9050

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	0.938424830	0.058670494	15.99483	0.00000000
2. APPNETS	-0.063680402	0.026823316	-2.37407	0.01759326
3. NETS	0.123664339	0.012356333	10.00817	0.00000000
4. APPLTV	-0.349407930	0.018079729	-19.32595	0.00000000
5. LTV	0.326047629	0.009384522	34.74313	0.00000000
6. LOGINCOME	-0.080710424	0.007175417	-11.24819	0.00000000

Probit Estimation for Home Loans Subsample

Binary Probit - Estimation by Newton-Raphson

Convergence in 4 Iterations. Final criterion was 0.0000000 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations	22368
Degrees of Freedom	22362
Skipped/Missing (from 28377)	6009
Log Likelihood	-14709.3273
Average Likelihood	0.5180902
Pseudo-R ²	0.0679331
Log Likelihood(Base)	-15476.4644
LR Test of Coefficients(5)	1534.2743
Significance Level of LR	0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	1.308572483	0.191792830	6.82284	0.00000000
2. APPNETS	-0.204264030	0.081581881	-2.50379	0.01228704
3. NETS	0.246565087	0.044733600	5.51185	0.00000004
4. APPLTV	-0.978222652	0.050592174	-19.33545	0.00000000
5. LTV	0.900451776	0.028196033	31.93541	0.00000000
6. LOGINCOME	-0.226599973	0.023488669	-9.64720	0.00000000

Logit Estimation for Home Loans Subsample

Binary Logit - Estimation by Newton-Raphson

Convergence in 4 Iterations. Final criterion was 0.0000006 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations	22368
Degrees of Freedom	22362
Skipped/Missing (from 28377)	6009
Log Likelihood	-14705.3348
Average Likelihood	0.5181827
Pseudo-R ²	0.0682832
Log Likelihood(Base)	-15476.4644
LR Test of Coefficients(5)	1542.2593
Significance Level of LR	0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	2.189339984	0.312397438	7.00819	0.00000000
2. APPNETS	-0.338278130	0.135727206	-2.49234	0.01269051
3. NETS	0.397687554	0.073469528	5.41296	0.00000006
4. APPLTV	-1.627625563	0.084803135	-19.19299	0.00000000
5. LTV	1.477352561	0.047459086	31.12897	0.00000000
6. LOGINCOME	-0.374498860	0.038337741	-9.76841	0.00000000

Linear-Probability Ordinary Least Squares Estimation for Home Loans Subsample

Linear Regression - Estimation by Least Squares

With Heteroscedasticity-Consistent (Eicker-White) Standard Errors

Dependent Variable DEFAULT

Usable Observations	22368
Degrees of Freedom	22362
Skipped/Missing (from 28377)	6009
Centered R ²	0.0661276
R-Bar ²	0.0659188
Uncentered R ²	0.5097671
Mean of Dependent Variable	0.4750536481
Std Error of Dependent Variable	0.4993884549
Standard Error of Estimate	0.4826483269
Sum of Squared Residuals	5209.2146491
Log Likelihood	-15441.4670
Durbin-Watson Statistic	1.9196

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	0.996323531	0.071732782	13.88938	0.00000000
2. APPNETS	-0.073444145	0.031389934	-2.33974	0.01929740
3. NETS	0.088995743	0.016291463	5.46272	0.00000005
4. APPLTV	-0.363894223	0.019491210	-18.66966	0.00000000
5. LTV	0.337626400	0.010234212	32.98997	0.00000000
6. LOGINCOME	-0.086025997	0.008792175	-9.78438	0.00000000

Probit Estimation for Buy-to-Let Loans Subsample

Binary Probit - Estimation by Newton-Raphson

Convergence in 4 Iterations. Final criterion was 0.0000016 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations	2625
Degrees of Freedom	2619
Skipped/Missing (from 28377)	25752
Log Likelihood	-1645.3332
Average Likelihood	0.5343022
Pseudo-R ²	0.1299221
Log Likelihood(Base)	-1819.1256
LR Test of Coefficients(5)	347.5848
Significance Level of LR	0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	0.897857031	0.431583136	2.08038	0.03749067
2. APPNETS	-0.054485092	0.146683025	-0.37145	0.71030400
3. NETS	0.677650230	0.076976891	8.80329	0.00000000
4. APPLTV	-0.682025821	0.149429489	-4.56420	0.00000501
5. LTV	0.749668838	0.073490300	10.20092	0.00000000
6. LOGINCOME	-0.218409769	0.048997293	-4.45759	0.00000829

Logit Estimation for Buy-to-Let Loans Subsample

Binary Logit - Estimation by Newton-Raphson

Convergence in 5 Iterations. Final criterion was 0.0000000 <= 0.0000100

Dependent Variable DEFAULT

Usable Observations	2625
Degrees of Freedom	2619
Skipped/Missing (from 28377)	25752
Log Likelihood	-1643.3048
Average Likelihood	0.5347153
Pseudo-R ²	0.1314084
Log Likelihood(Base)	-1819.1256
LR Test of Coefficients(5)	351.6416
Significance Level of LR	0.0000000

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	1.831993788	0.751513100	2.43774	0.01477939
2. APPNETS	-0.086471754	0.243180497	-0.35559	0.72215009
3. NETS	1.115593393	0.131803868	8.46404	0.00000000
4. APPLTV	-1.124528032	0.247846536	-4.53719	0.00000570
5. LTV	1.234899192	0.122897219	10.04823	0.00000000
6. LOGINCOME	-0.401753890	0.086007032	-4.67117	0.00000299

Linear-Probability Ordinary Least Squares Estimation for Buy-to-Let Loans Subsample

Linear Regression - Estimation by Least Squares

With Heteroscedasticity-Consistent (Eicker-White) Standard Errors

Dependent Variable DEFAULT

Usable Observations	2625
Degrees of Freedom	2619
Skipped/Missing (from 28377)	25752
Centered R ²	0.1220452
R-Bar ²	0.1203691
Uncentered R ²	0.5534973
Mean of Dependent Variable	0.4914285714
Std Error of Dependent Variable	0.5000217765
Standard Error of Estimate	0.4689636333
Sum of Squared Residuals	575.98852326
Log Likelihood	-1733.9813
Durbin-Watson Statistic	1.9313

Variable	Coefficient	Std Error	T-Stat	Significance
1. Constant	0.741288371	0.181985364	4.07334	0.00004634
2. APPNETS	-0.021689829	0.052459785	-0.41346	0.67927233
3. NETS	0.242973400	0.026289725	9.24214	0.00000000
4. APPLTV	-0.239099705	0.053247424	-4.49035	0.00000711
5. LTV	0.263310913	0.024072584	10.93821	0.00000000
6. LOGINCOME	-0.068559400	0.020681416	-3.31502	0.00091635

Next, we examine residual plots for the ordinary least squares model, keeping in mind that heteroskedasticity is a necessary component of this model when the endogenous variable is binary, as in our case. We plot the residuals for each of the three regression-based estimates (all, home, and buy-to-let subsamples) against each of the five explanatory variables. These are shown on the following fifteen graphs (Figures A.1.1 – A.1.15). Since the endogenous variable is binary and the explanatory variables are not, this model always has conditional heteroskedasticity by construction (see Greene (2008, pp. 772-773)) but other than this effect there are no obvious or extreme heteroskedasticity effects detectable.

Figure A.1.1

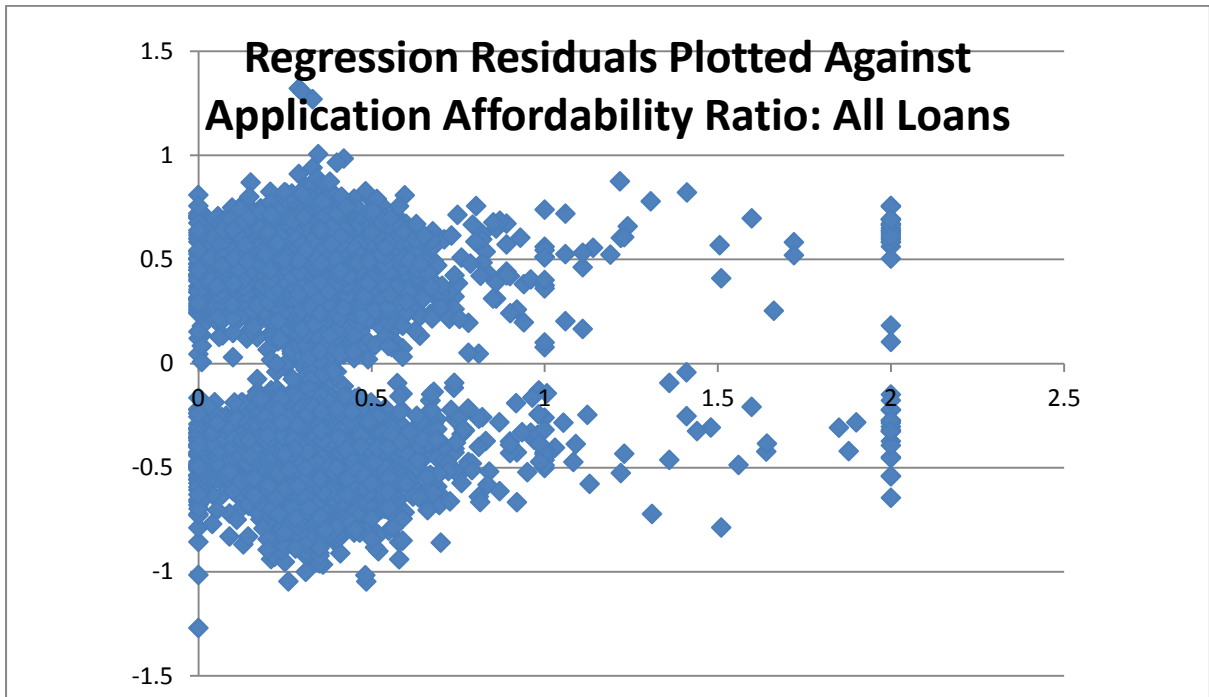


Figure A.1.2

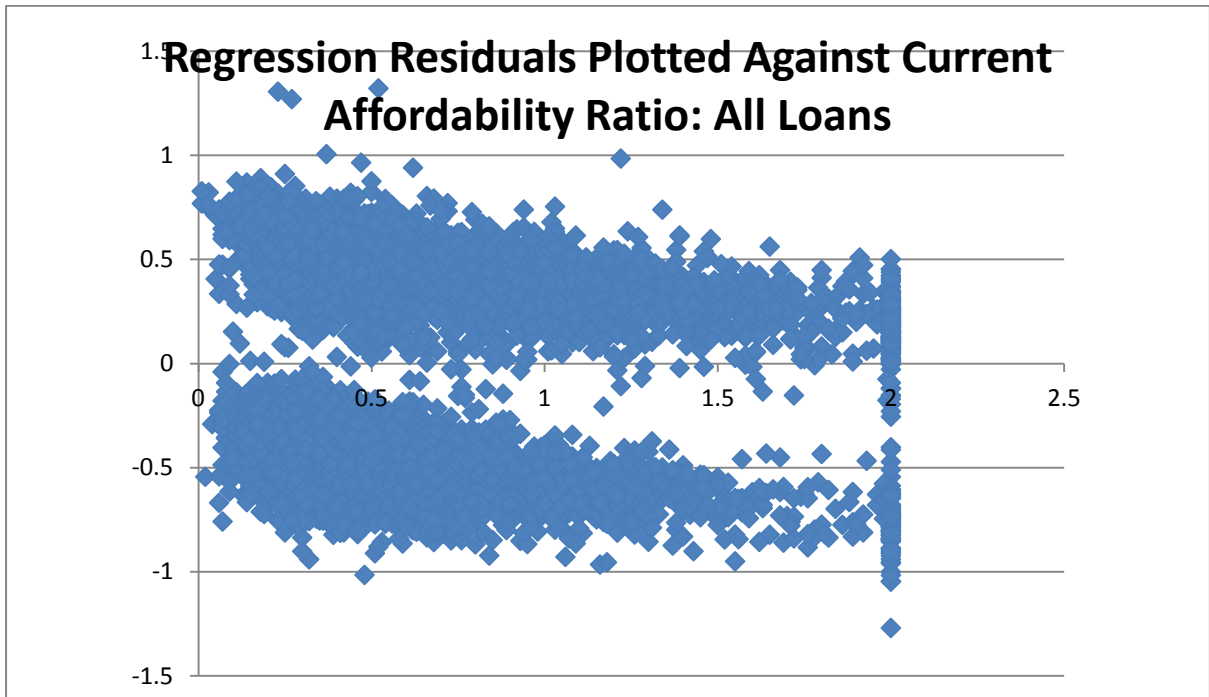


Figure A.1.3

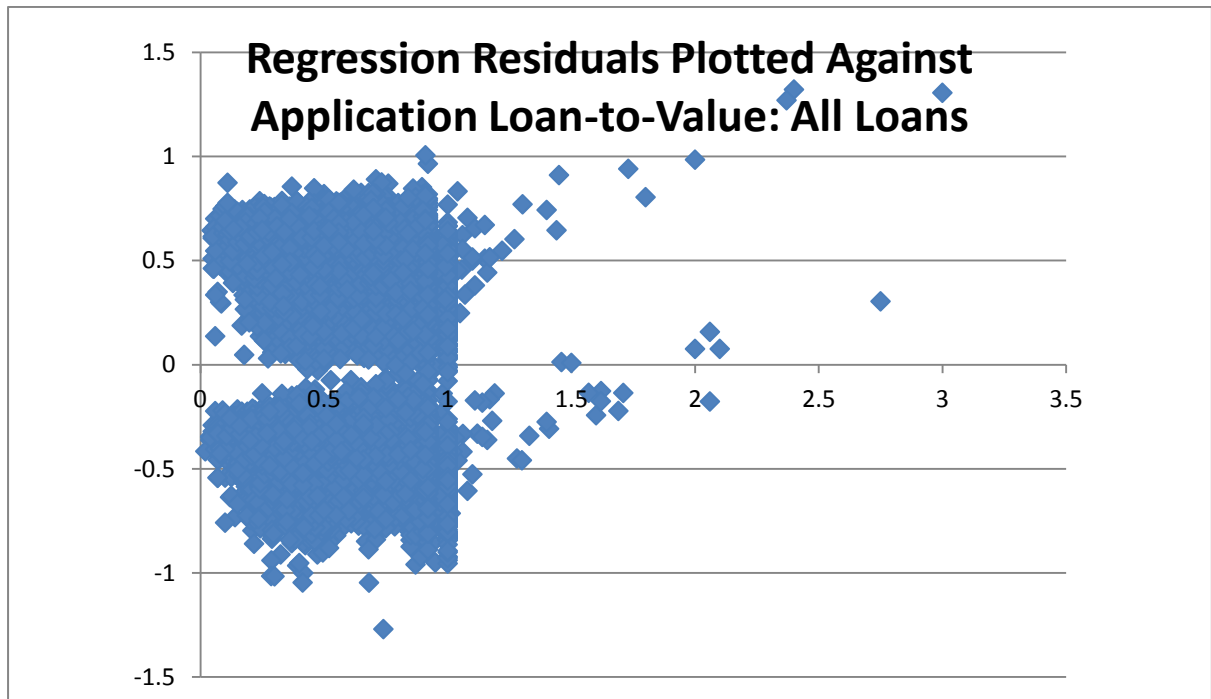


Figure A.1.4

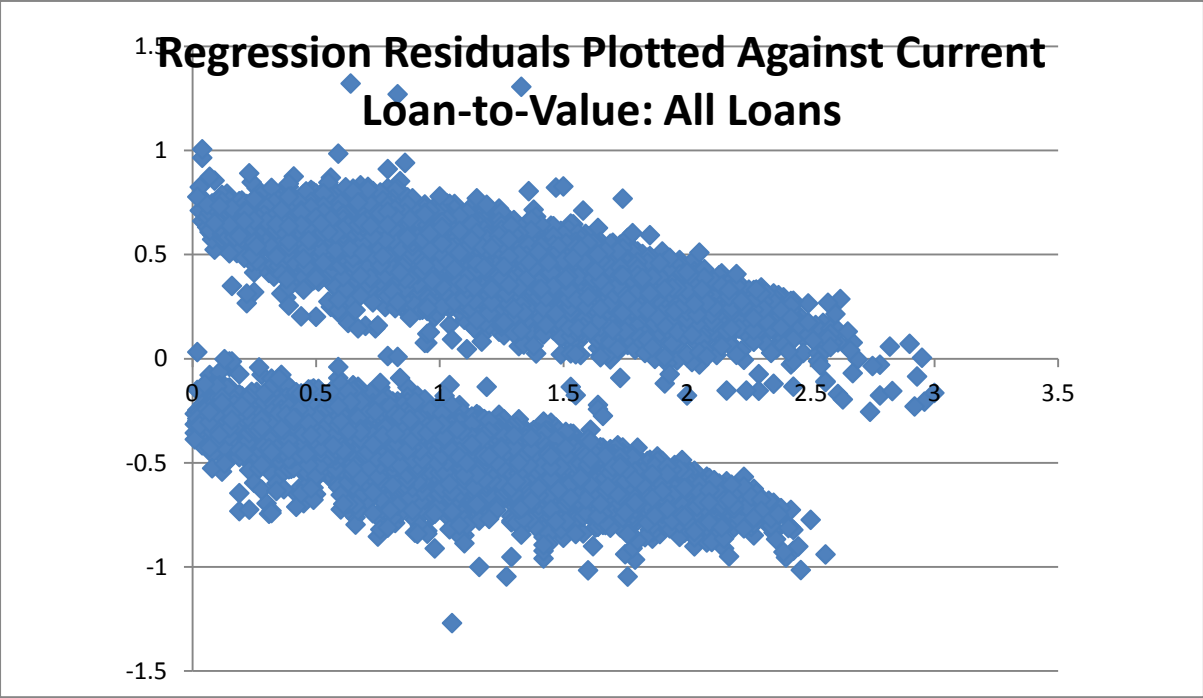


Figure A.1.5

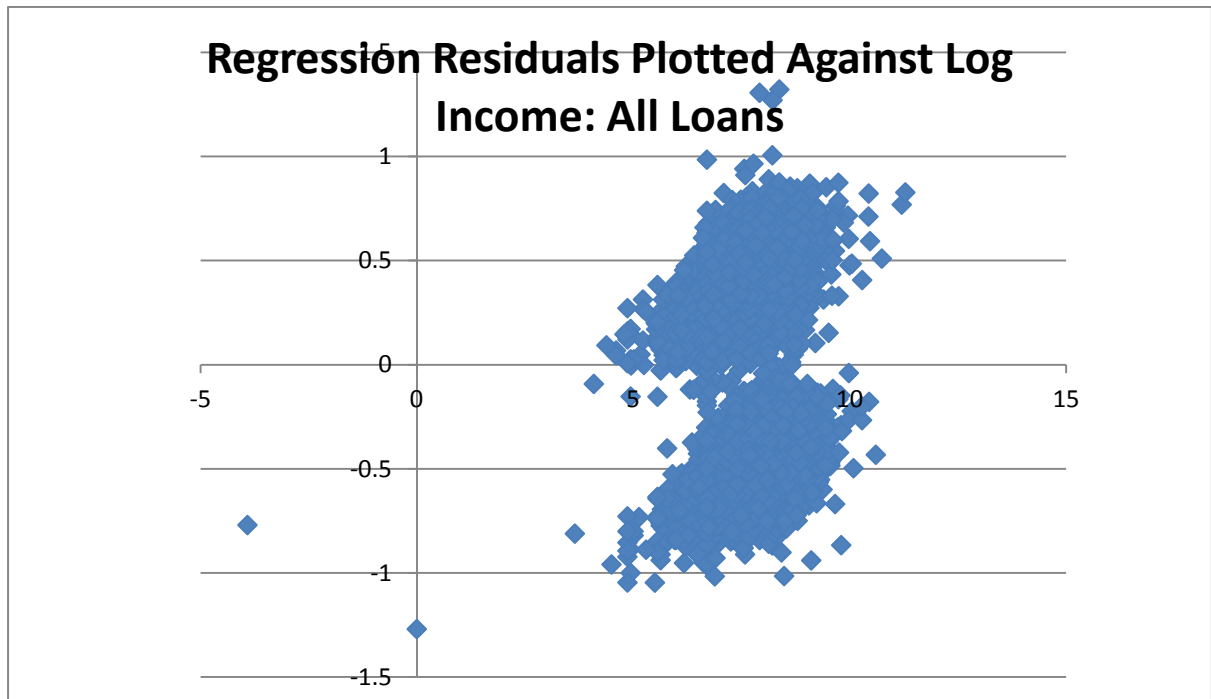


Figure A.1.6

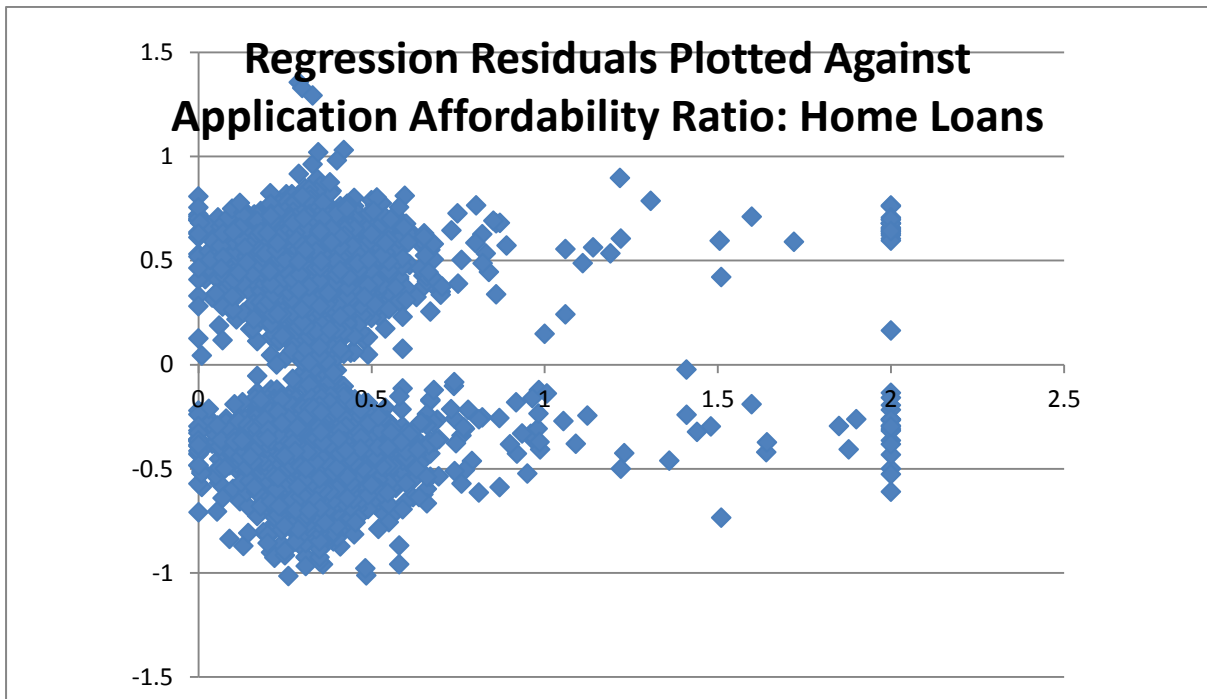


Figure A.1.7

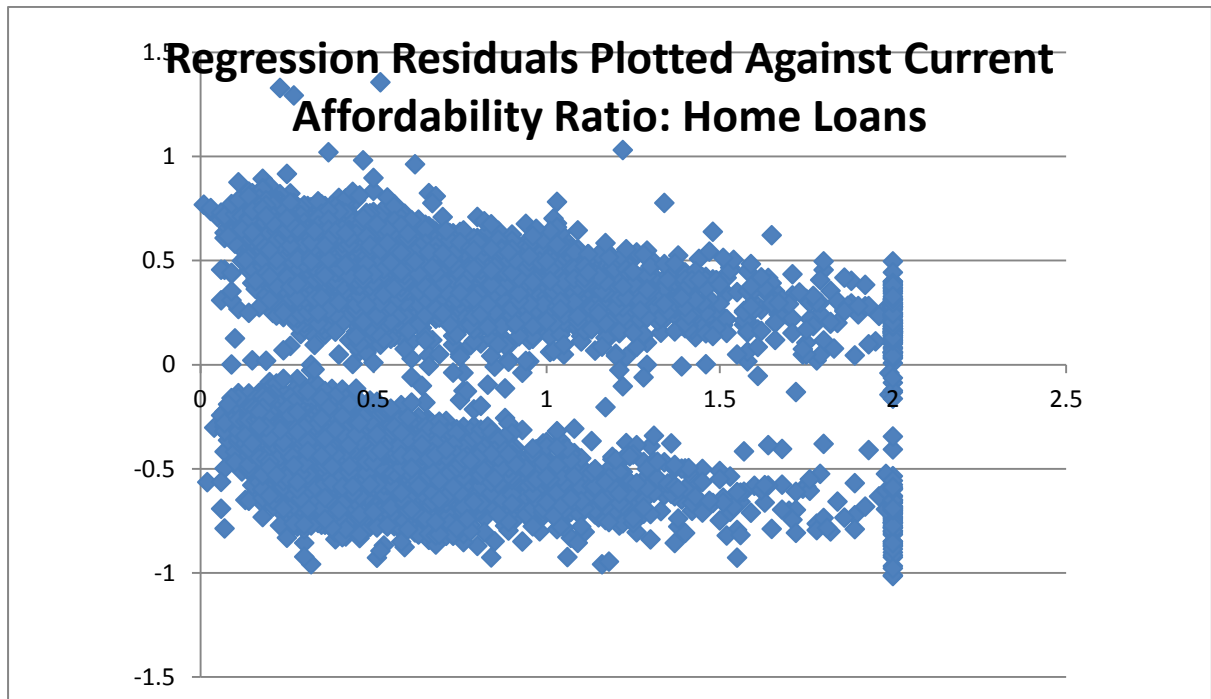


Figure A.1.8

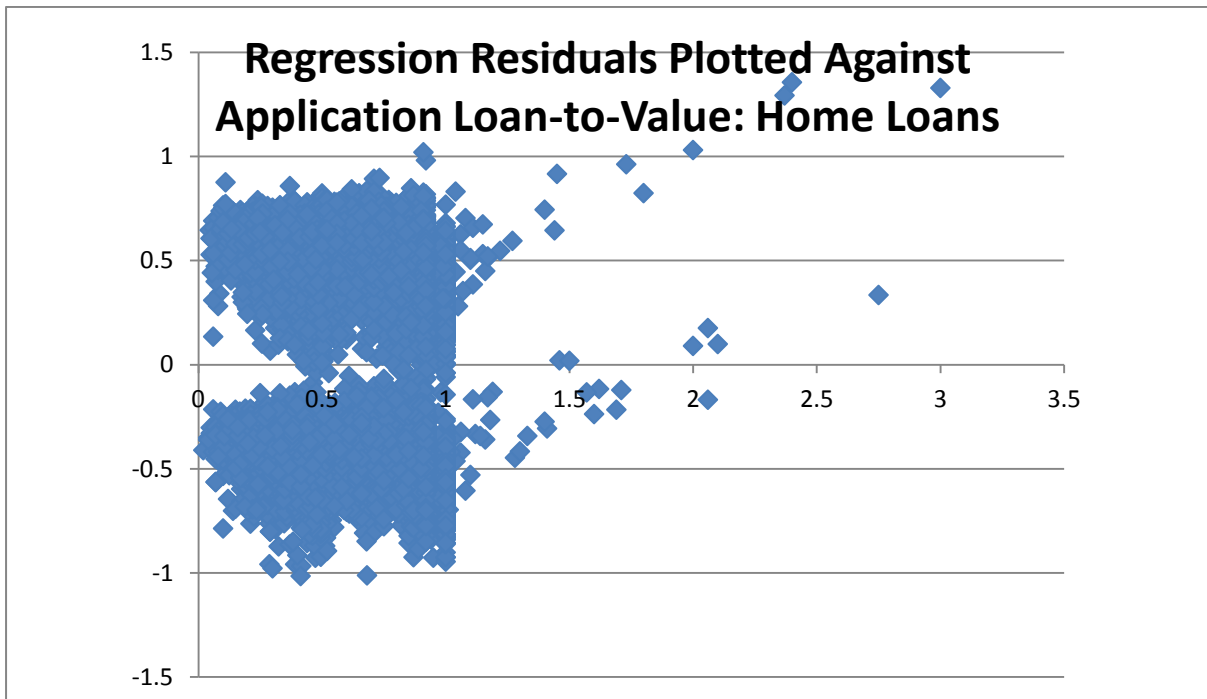


Figure A.1.9

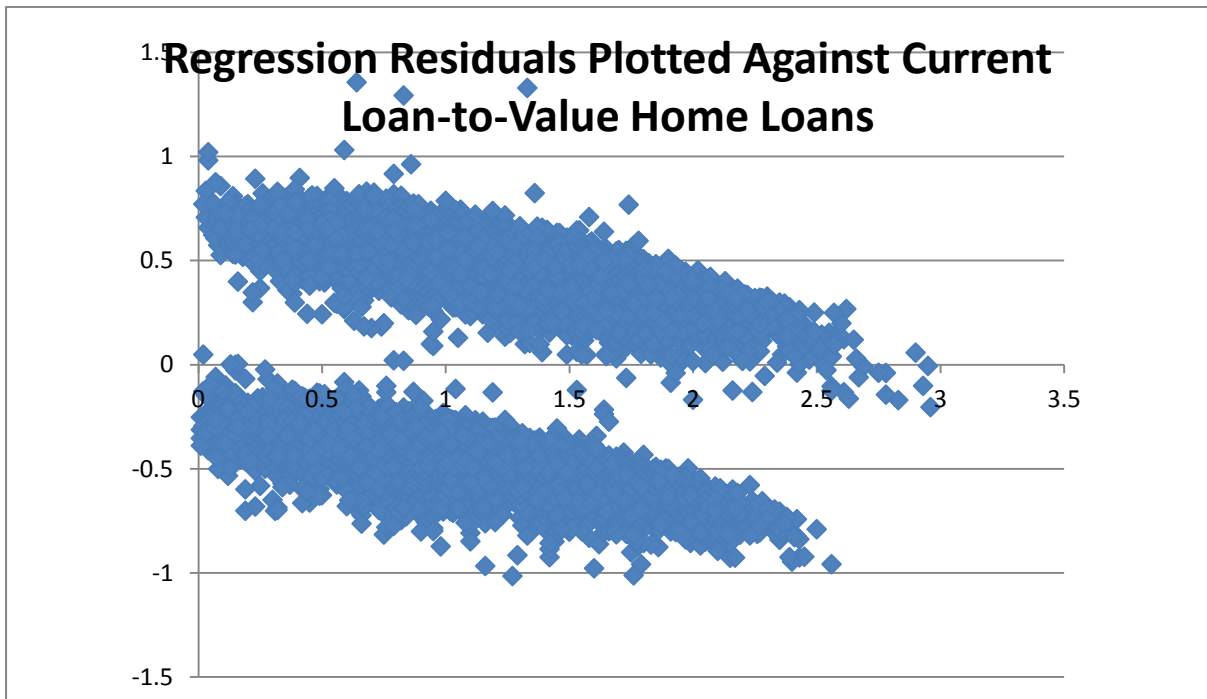


Figure A.1.10

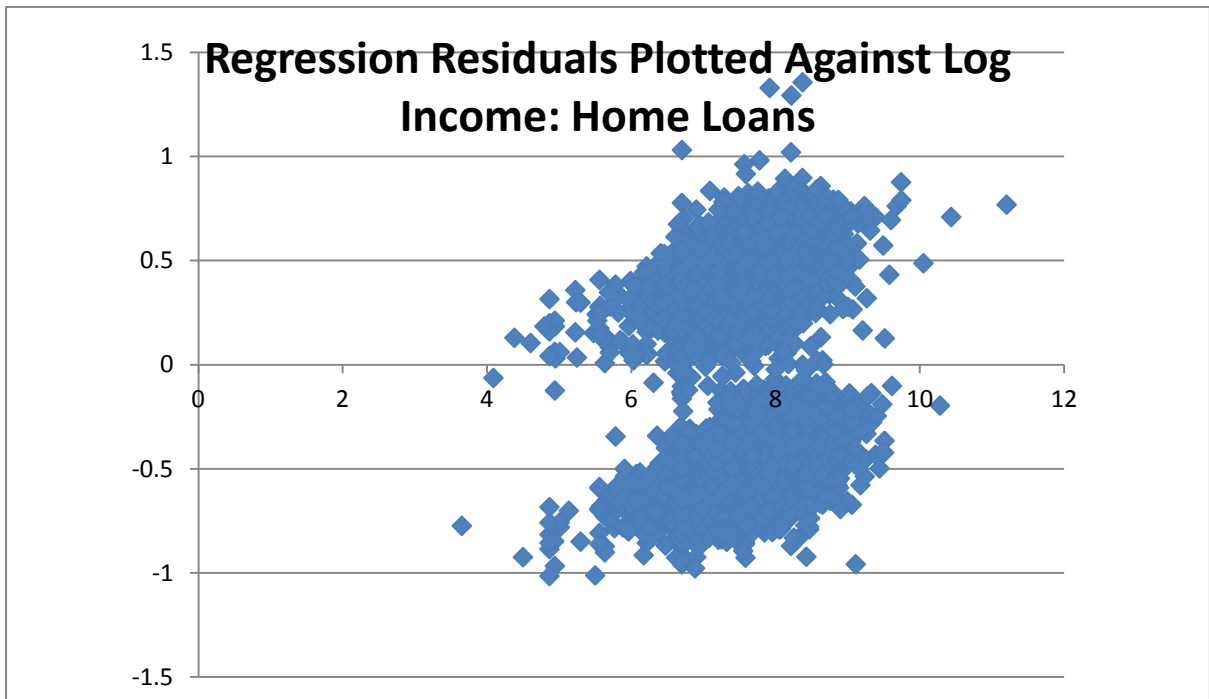


Figure A.1.11

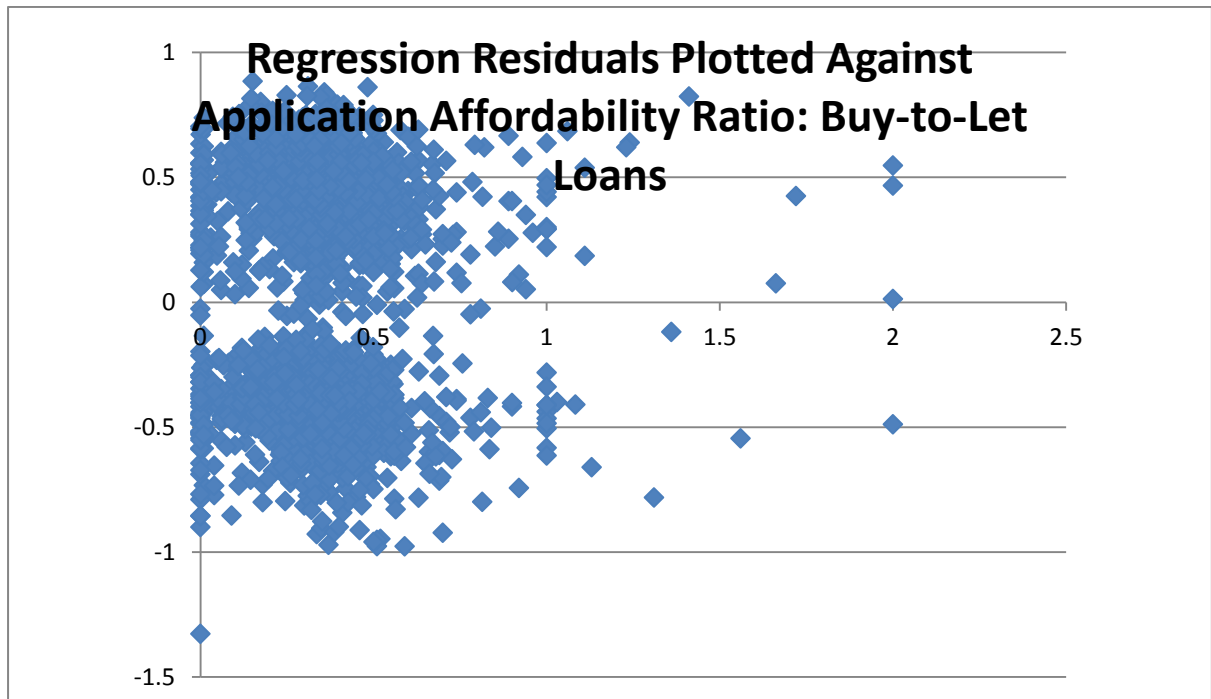


Figure A.1.12

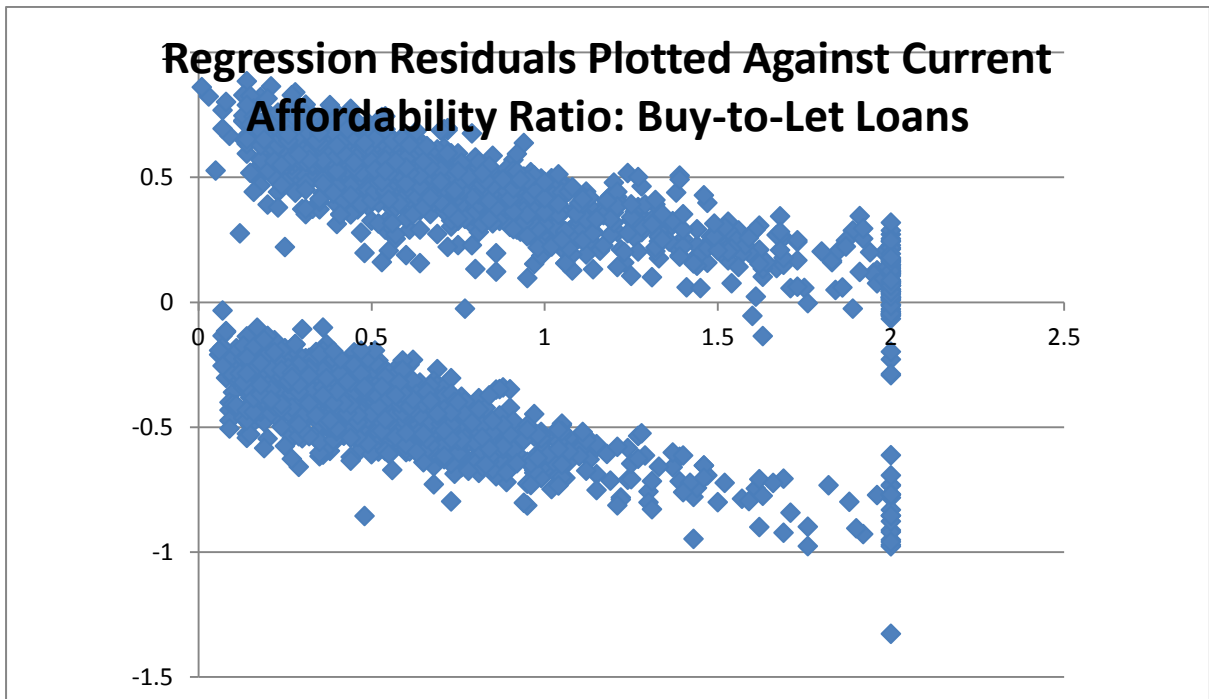


Figure A.1.13

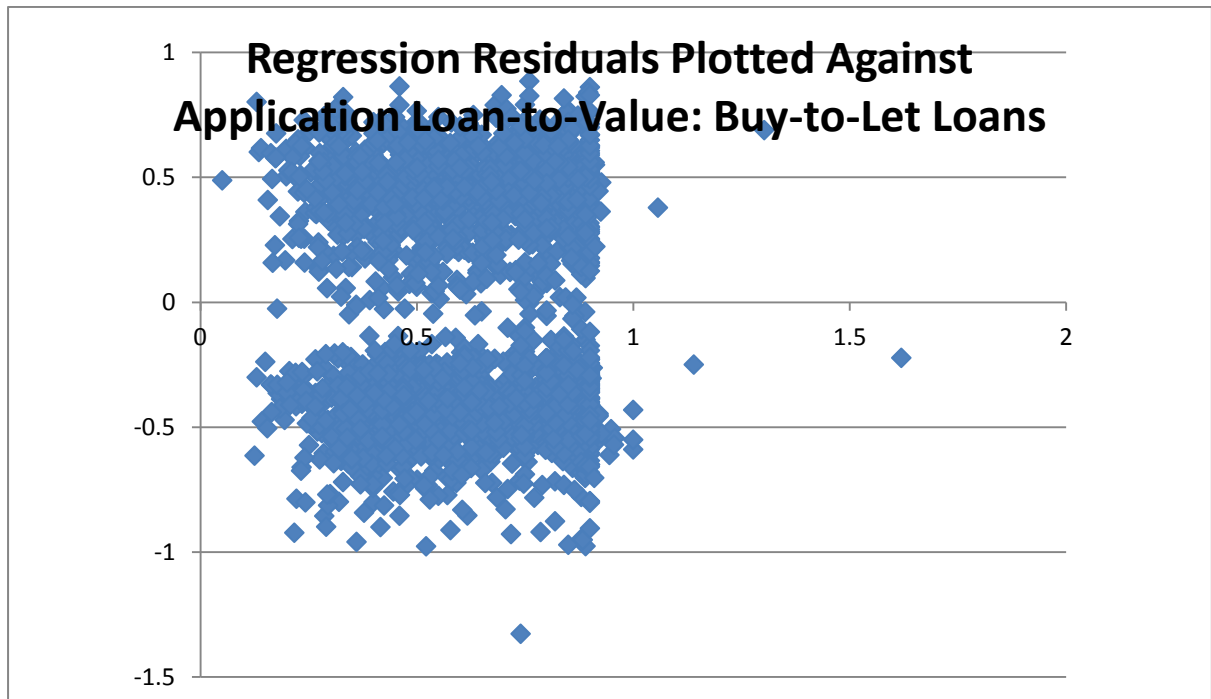


Figure A.1.14

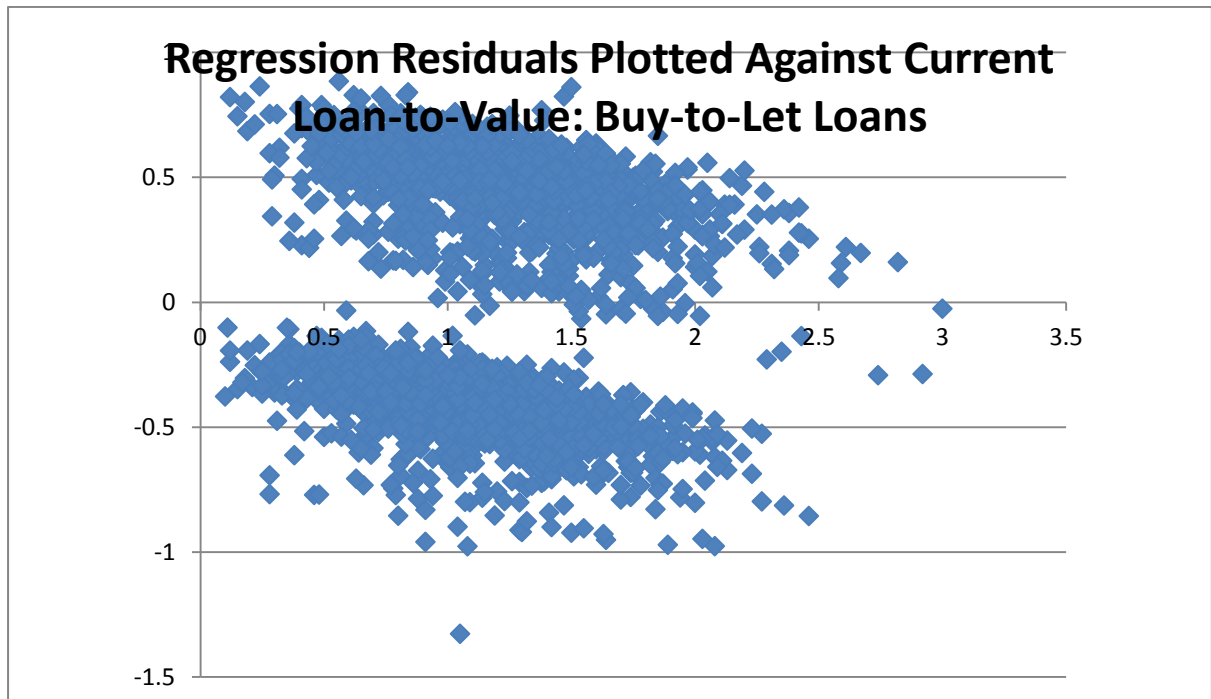
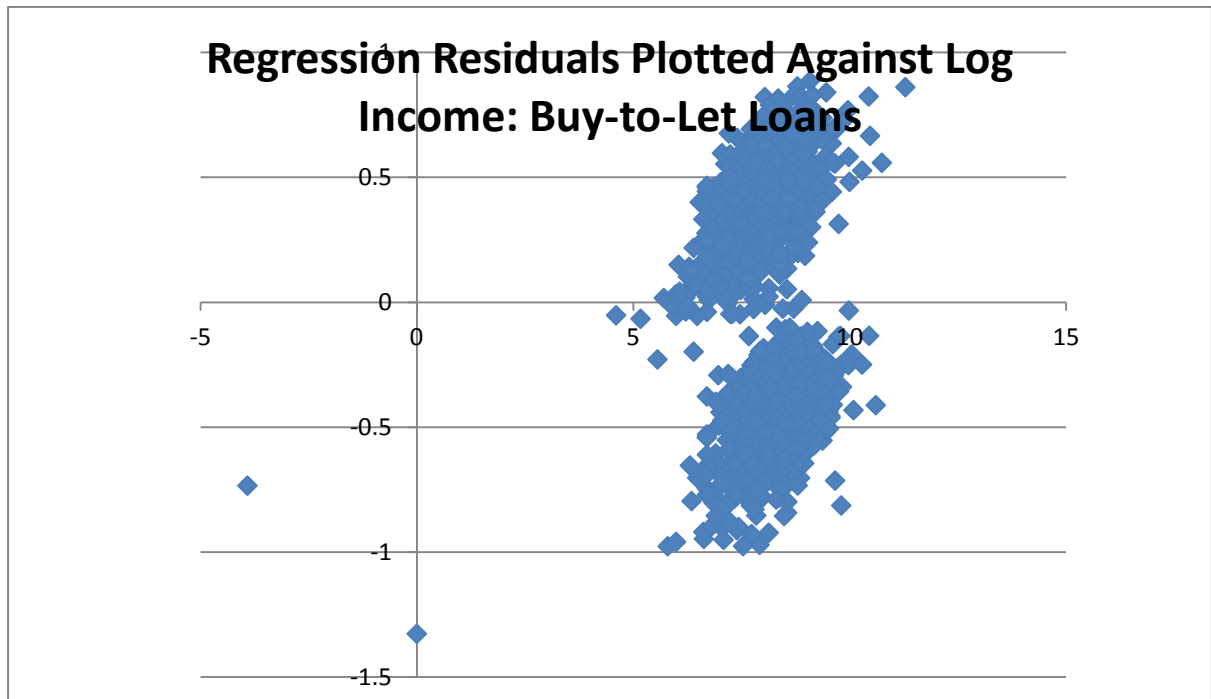


Figure A.1.15



2. Ancillary Tables

This section presents some extra tables which are mentioned in the paper but not of sufficient importance to include in the main paper.

Table A.2.1: Default rates for loans doubly-sorted by loan-to-value, payment-to-income and net income: Home loans

A.2.1.a: Default rates for loans sorted by loan-to-value and affordability

	Low Payment-to-Income	Moderate Payment-to-Income	High Payment-to-Income
Low LTV	43.70%	39.44%	45.77%
Moderate LTV	42.70%	44.57%	56.97%
High LTV	48.40%	56.84%	70.69%

A.2.1.b: Default rates for loans sorted by loan-to-value and net income

	High Income	Moderate Income	Low Income
Low LTV	32.70%	40.11%	48.63%
Moderate LTV	38.67%	46.83%	55.69%
High LTV	51.08%	59.41%	67.97%

A.2.1.c: Default rates for loans sorted by affordability and net income

	High Income	Moderate Income	Low Income
Low Payment-to-Income	39.37%	43.08%	45.69%
Moderate Payment-to-Income	45.28%	47.43%	59.59%
High Payment-to-Income	58.59%	48.32%	59.66%

Table A.2.2: Default rates for loans doubly-sorted by loan-to-value, payment-to-income and net income: Buy-to-let loans

A.2.2.a: Default rates for loans sorted by loan-to-value and affordability

	Low Payment-to-Income	Moderate Payment-to-Income	High Payment-to-Income
Low LTV	33.78%	31.21%	47.89%
Moderate LTV	27.95%	37.50%	58.35%
High LTV	39.10%	51.53%	70.04%

A.2.2.b: Default rates for loans sorted by loan-to-value and net income

	High Income	Moderate Income	Low Income
Low LTV	27.49%	43.90%	57.14%
Moderate LTV	40.29%	53.44%	69.85%
High LTV	52.15%	64.78%	85.00%

A.2.2.c: Default rates for loans sorted by affordability and net income

	High Income	Moderate Income	Low Income
Low Payment-to-Income	31.44%	38.29%	53.08%
Moderate Payment-to-Income	36.07%	43.94%	62.06%
High Payment-to-Income	50.00%	51.85%	73.86%

Table A.2.3: Partially-linear Index Probit Model of Default with Four Explanatory Variables

Variable	Full Sample		Home Loans		Buy-to-Let Loans	
	Estimated Coefficient with Standard Probit	Estimated Coefficient with Partially Linear Index Probit	Estimated Coefficient with Standard Probit	Estimated Coefficient with Partially Linear Index Probit	Estimated Coefficient with Standard Probit	Estimated Coefficient with Partially Linear Index Probit
App LTV	-0.394	-0.471	-0.383	-0.466	-0.425	-0.434
Affordability	0.227	0.227	0.093	0.086	0.491	0.497
LTV	0.452	Nonparametric	0.458	Nonparametric	0.483	Nonparametric
Log Income	-0.237	-0.241	-0.275	-0.283	-0.210	-0.217
Constant	1.484	N/A	1.825	N/A	1.089	N/A

3. Successful Workouts and Cures

In many jurisdictions, a “cure” rate analysis is critically important. However, this type of analysis is not very relevant / informative in the Irish case over our sample time period. During our sample period, the Irish system of mortgage workout was in near paralysis rather than having any dynamic process of cure or repossession following default. For example, in our sample, computing a cure as a loan that was in default on SFS date and performing at the sample end (August 2013) would leave us with an extremely small sample. Figure A.3.1 shows the proportion of loans in default on SFS date that resulted in cures. Figure A.3.2 shows the cumulative number of cures post-SFS date. The number of cures achieved from the early SFS submissions is extremely small; the number increases rapidly, but still leaves us with only 504 cures. Furthermore, caution must be exercised in interpreting these workouts as true “cures” because some of them might be single payments or a small number of payments during the negotiation process followed by a quick relapse into the ‘default’ category. Though they satisfy our definition of a cure at the sample end, we have no evidence (for or against) that these loans continued to perform.

We do not have any information on repossessions for our sample of loans, though we know from informal discussions with the data providers that it is very low. In the table below we rely on data from the Central Bank of Ireland to depict the trend in repossessions across all Irish mortgage providers, including but not restricted to our data provider. Figure A.3.3 shows the number of repossessions as a proportion of all residential mortgages that are in default (accumulated arrears greater than 90 days worth of payments). The repossession rate is very low, always less than 1.5% and actually declines over time, i.e. the number of mortgages in default grows more quickly than the number of repossessions, so that the rate falls rather than increases over time. For these reasons, we do not undertake “cure” analysis in the paper.

Figure A.3.1

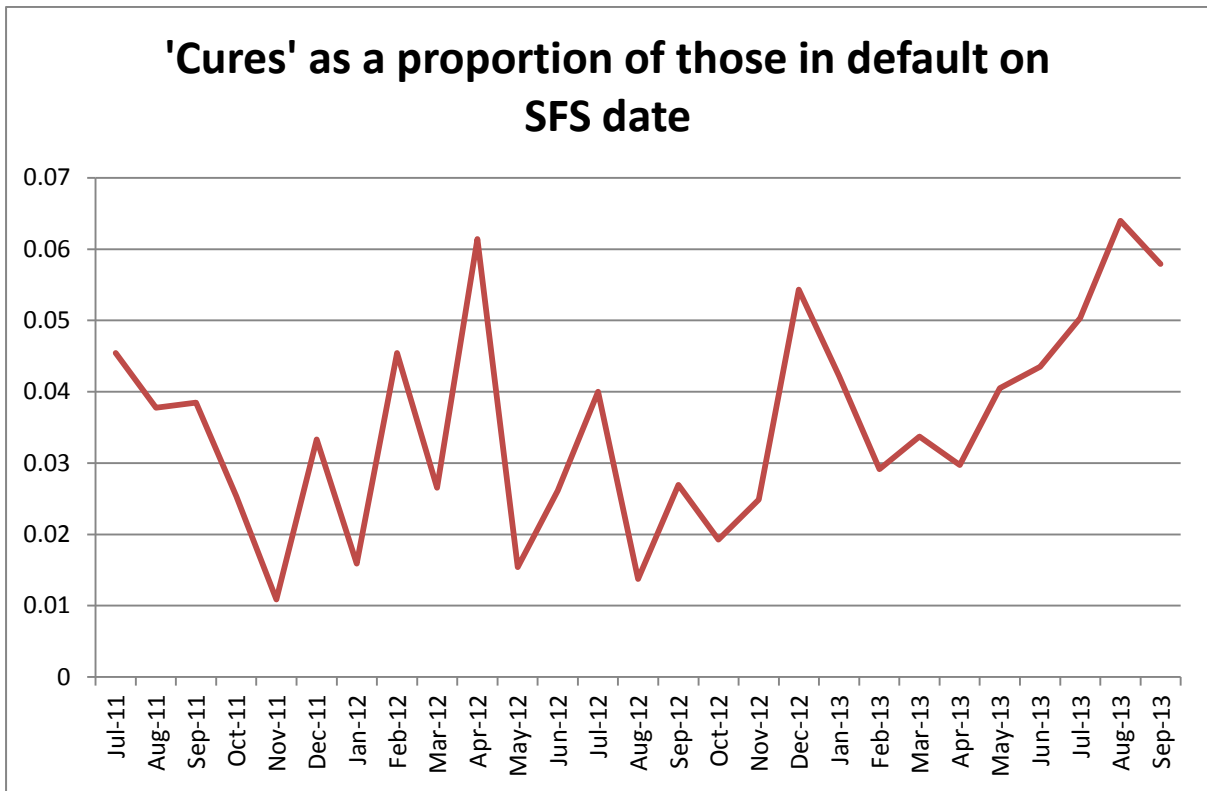


Figure A.3.2

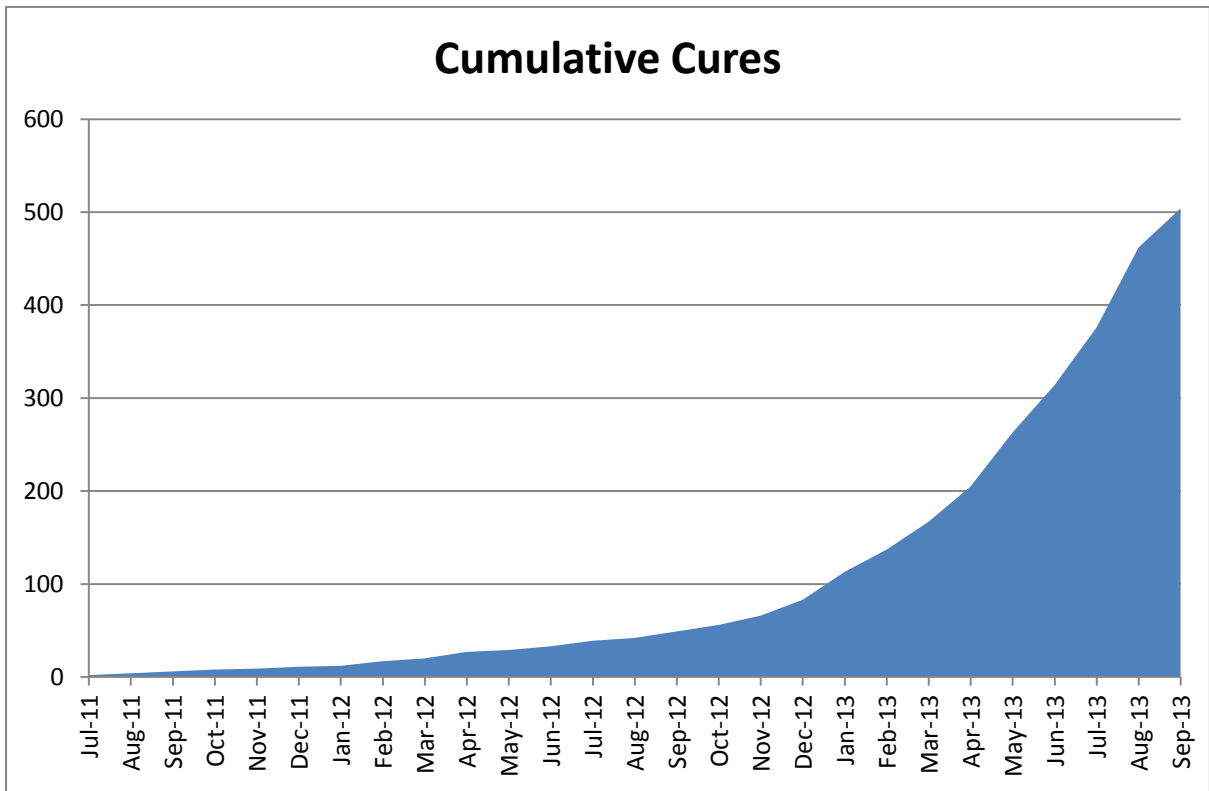
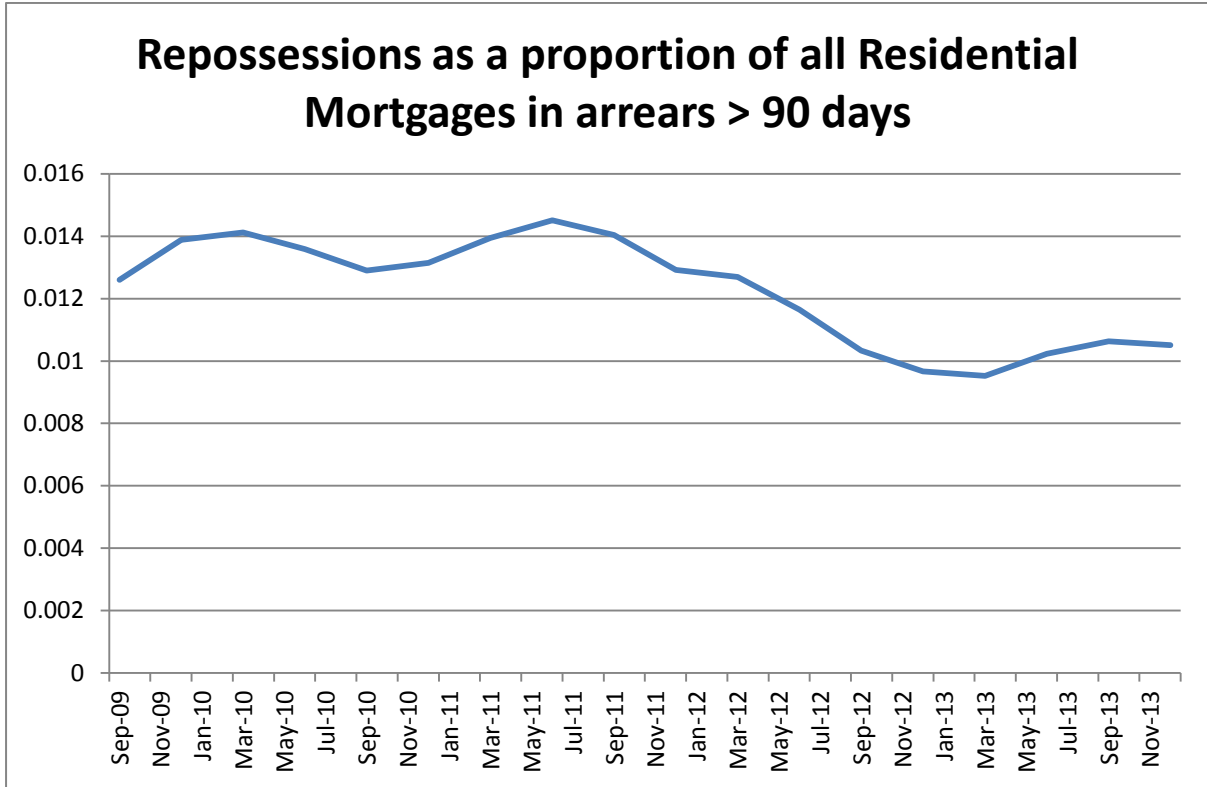


Figure A.3.3



Bibliography

Connor, Gregory and Thomas Flavin (2014) "Strategic, Unaffordability, and Dual-Trigger Default in the Irish Mortgage Market," working paper, National University of Ireland, Maynooth.

Enders, Walt, (2004) *RATS Programming Manual*, published by Estima Inc., e-book available at <http://www.estima.com/enders/>.

Greene, William H., (2008), *Econometric Analysis*, Sixth Edition, Prentice-Hall Inc.