Mortgage Loan Approvals and Government Intervention Policy

Dr. William Chow 18 March, 2014

Executive Summary

- This paper introduces an empirical framework to explore the impact of the government's various cooling measures on banks' mortgage lending behavior.
- Loan to value ratio and contractual life are among our explanatory variables. While the government can intervene in the market by imposing restrictions on lending through these channels, we use an indicator variable to differentiate the pre and post intervention periods. Incorporating interactive terms in our model provides an assessment of the incremental effect of the explanatory variables (in addition to their regular marginal effects) due to the presence of government intervention.
- The overall evidence suggests that the responses of loan approval are indeed different with and without the intervention measures, although the significance varies when it comes to the interaction with individual explanatory variables.
- It seems that banks are themselves risk averse towards mortgage lending. The amount of loan applications rejected increases with leverage (the LTV ratio), market activities (S&P), the share of primary market deals and reduction in co-financing deals. They also tend to reject more often in low interest rate environment, but are not particularly responsive to property price changes.

1. Introduction

1.1 We analyze in this paper how the government intervention measures in the property market influence mortgage lending behavior of banks. As our data run from 2003 onwards, the interventions recorded in the sample period cover only the various "cooling" measures adopted by the Hong Kong government from 2009 to 2013. Instead of direct causality, we modeled the impact on banks via interactions between the measures and various market and bank related variables.

1.2 With the benefit of hindsight, the studied relationship should be understood as complementary rather than contradictory as property transactions and the amount of loan deals dropped significantly after the adoption of the measures. Still, our analysis provides insights into how the banks would have behaved had the measures not been implemented.

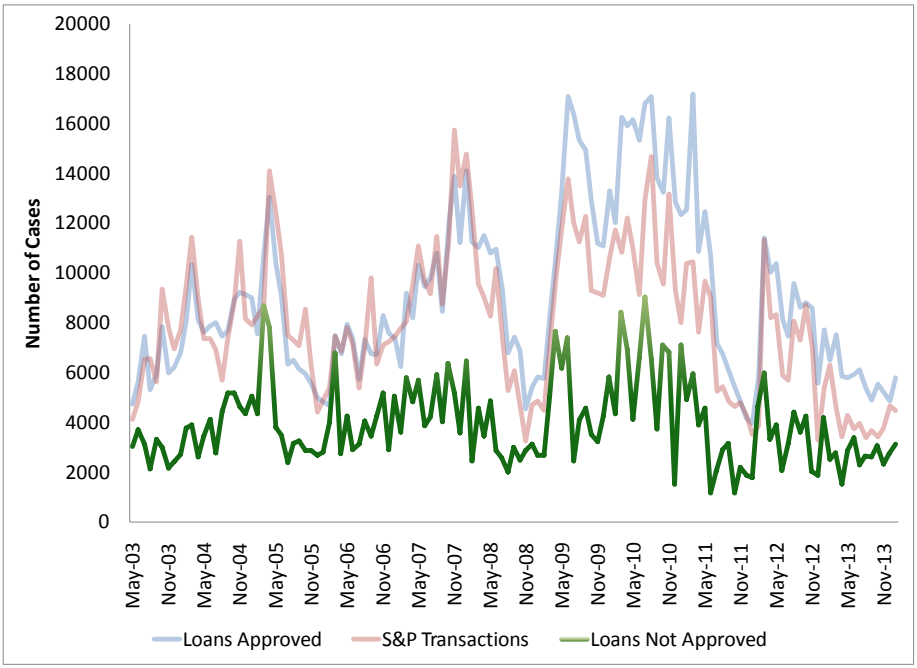
1.3 The approach we used is the regression of count data; see for example the book of Cameron and Trivedi (1998) or Winkelmann (2008). Count data involve discrete variables such as number of occurrences of events and cannot be evaluated using ordinary least squares method. For reasons to be explained shortly, we used negative binomial regression to conduct the analysis. The case we considered is univariate in nature, but it can be extended to multivariate formulations at the expense of higher model and estimation complexity.

1.4 The relationship between various bank (and market) variables and mortgage lending behavior turns out to be regime specific. That is, how the former affect the latter would depend on the intervention policy of the government. The report is organized as follows: Section 2 discusses the data we used and the intuition of the model choice; Section 3 highlights the modeling basics; Section 4 summarizes the estimation output and the findings; and Section 5 concludes.

2. The Data

2.1 The source of the data is Hong Kong Monetary Authority's monthly residential mortgage survey. The usable sample[1] runs from May, 2003 to January, 2014. The principal variable is the number of unapproved mortgage applications. We may as well use the number of approved cases but did not do so because it is highly correlated with property transactions and has too much persistence (hence, less evidence of independent observations). Figure 1 shows the Loans Approved (LA), Loans Unapproved (LU) and the Sales & Purchase transaction (S&P) series, all in number of cases.

Figure 1. S&P Transactions and Mortgage Loans

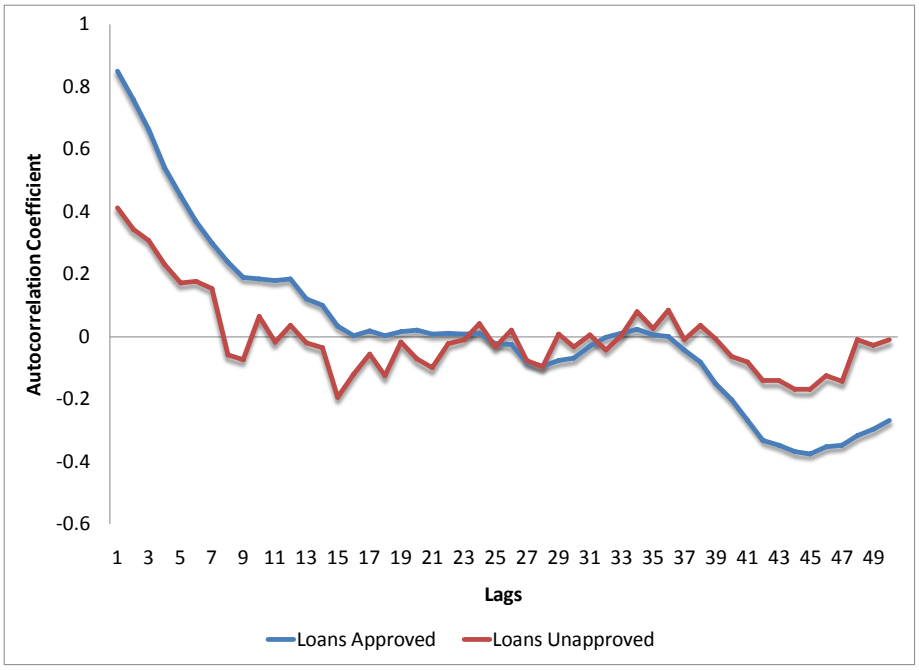


2.2 One can see that S&P is more highly correlated with LA than with LU. Since S&P is not modeled endogenously in this paper, we run the risk of ignoring or misinterpreting feedbacks between S&P and LA if the latter is used as the independent variable. A second reason for

[1] The survey started in 1997 but the format of the survey changed somewhat over time. After incorporating needed data transformations, we take the present sample which runs from May, 2003 to January, 2014.

choosing LU is that there is less persistence in the data. Our regression analysis requires the data to be independent observations. Figure 2 shows that LA has higher autocorrelations at all lags up to 22 than LU, so LU is in a better position to meet the model requirement.

Figure 2. Autocorrelation at Different Lag Levels for LA and LU



2.3 The other variables involved in the econometric exercise are: (i) Loan to Value ratio (LTV), (ii) contractual life of the mortgage (CL), (iii) year-on-year change in residential property price index (ΔP), (iv) S&P, (v) 3 month HIBOR (HIBOR-3), (vi) share of approved loans with co-financing (COF), (vii) share of approved loans with primary market transactions (PRIM), (viii) value of loans written off in the past 12 months (WOFF), and (ix) a concentration measure of financing structure (HHI).

2.4 Items (i) to (viii) can be found in the mortgage surveys and from official statistical releases. Item (ix) is an indicator showing the financing interest rate spectrum of the mortgagors. It is modeled using the Herfindahl-Hirschman Index:

$$HHI = \sum_{j=1}^{J} s_j^2 \qquad (1)$$

where s_j is the share of loans arranged under terms as specified by plan j. The $J = 4$ schemes we considered are (a) best lending rate (BLR) based mortgage rate below 2.5%, (b) BLR based mortgage rate between 2.5% and 3%, (c) BLR based mortgage rate in excess of 3%, and (d) HIBOR based mortgage rates. The larger the value of HHI, the financing term is skewed more to (concentrated on) a particular scheme.

2.5 During the earlier years of the mortgage survey, the financing terms of BLR based schemes were reported as margins above and below the BLR. From late 2009 onwards, they are reported as actual interest rate levels. To compile the HHI, we have to reconcile the reported data by converting all premia and discounts to BLR into actual interest rate terms. This is done by referring to the monthly BLR levels, simulating artificial contractual interest rates and regrouping the shares defined by the designated thresholds described in the paragraph above. Figure 3 plots the HHI together with the 3 month HIBOR. In general, mortgaging financing terms are more concentrated in low interest rate environment and vice versa.

2.6 Finally we plot the main dependent variable LU in relation to variables (i) – (viii) in Figures 4 and 5. The scatter diagrams split the data into two parts, the first being data before October, 2009 and the second those after that date. The date signifies the first important move by the government to cool off the market by way of a LTV upper bound of 60% for properties worth more than HK$20 m.

Figure 3. Herfindahl-Hirschman Index and 3 Month HIBOR

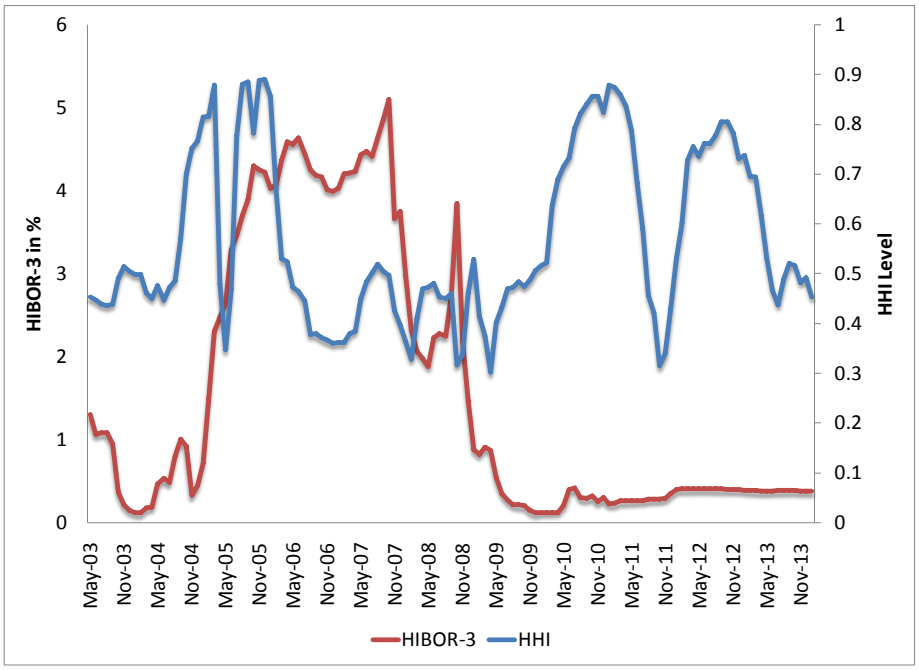


Figure 4. Loans Unapproved with LTV ratio, Contractual Life, Property Price Change and S&P

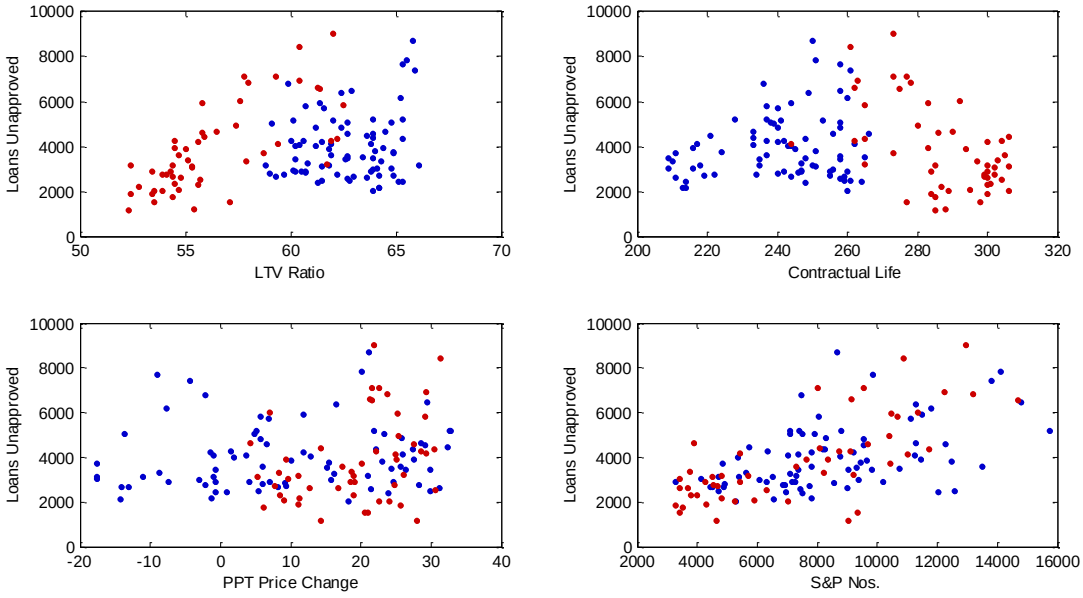
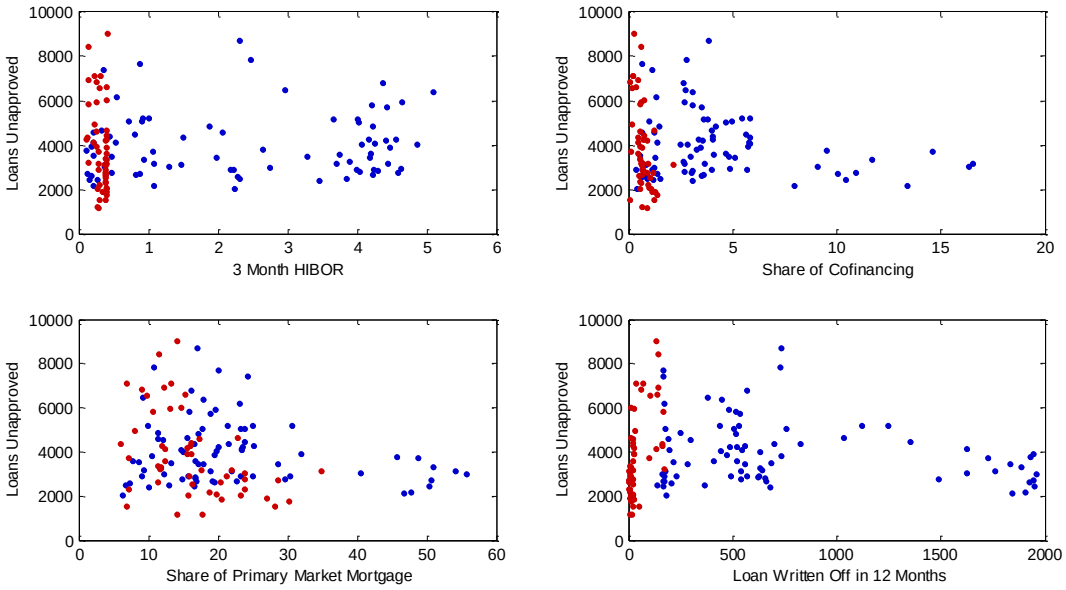


Figure 5. Loans Unapproved with HIBOR-3, Co-financing Share, Share of Primary Market Mortgage and Loans Written Off



2.7 Other important dates include:

- February, 2010 – stamp duty raised to 4.25% for properties worth more than HK$20 m.
- August, 2010 – confirmor transactions for pre-sale units banned; bank stress tests; LTV ratio of 60% for properties worth HK$12 m. or more.
- October-November, 2010 – My Home Purchase Plan; Special Stamp Duty (SSD) for all properties.
- June, 2011 – lower LTV for overseas buyers; for properties worth between HK$7-10 m. and those worth between HK$10-12 m., the LTV is lowered to 60% and 50% respectively; maximum loan amount of HK$4.2 m. for properties worth less than HK$7 m.

- September, 2012 – maximum 30 years mortgage; mortgage payment on property for investment purpose cannot account for more than 40% of income.
- October, 2012 – Buyer Stamp Duty of 15% for non permanent residents and foreign companies; SSD for another 3 years.
- February, 2013 – Double Stamp Duty of as much as 8.5% for deals in excess of HK$2 m.

2.8 From Figure 4 and 5, it is obvious that the correlation of some of the highlighted variables with LU behave differently before and after the threshold date. LTV, CL, HIBOR-3, COF and WOFF all show two clusters when plotted against LU[2]. The LU-LTV clusters appear to have the same sign but with a parallel displacement. The LU-CL clusters appear to have different signs (hence, correlation). The patterns of the clusters for LU-HIBOR-3, LU-COF and LU-WOFF are all distinctly different.

3. The Modeling Approach

3.1 As commented before, both LA and LU are count data and require special regression models to perform the statistical analysis[3]. Simple and natural candidates are Poisson regression (POIREG) and Negative binomial regression (NB2REG). They will be discussed in this section. Another issue is to find an approach that can accommodate the *structural break* observed in the data.

3.2 Our model choice is NB2REG and it is closely related to POIREG. The presence of the structural break will be modeled via interactive terms involving the explanatory factors and a dummy variable represented

[2] Interest rate, the amount of loans involving co-financing, and the amount of loans written off could be exogenous and follow developments not directly related to or influenced by government interventions.

[3] An alternative is to model the approval rate (the ratio of approved loans to total applications) in which case the dependent variable will be continuous and classical linear regression will probably suffice.

the periods under property market interventions. To facilitate understanding of the core model, introductory knowledge of the POIREG will be helpful.

3.3 Poisson regression can be specified by:

$$Prob(Y = y_i|\boldsymbol{X}_i) = \frac{e^{-\lambda_i}{\lambda_i}^{y_i}}{y_i!}, \quad (2)$$

$$E(y_i|\boldsymbol{X}_i) = Var(y_i|\boldsymbol{X}_i) = \lambda_i = e^{(\alpha+\boldsymbol{X}_i'\beta)}, \quad (3)$$

$$y_i = 0,1,2,\cdots; \quad i = 1,\cdots,N.$$

where y_i is the observed count for observation (period or individual) i; $\boldsymbol{X}_i$ is the vector of explanatory variables; and λ_i is both the conditional mean and variance of the dependent variable. The regressors enter the exponential in equation (3) linearly with an intercept α and coefficient vector β that help define the mean and variance of the count variable. The conditional mean, in turn, defines the probability of occurrences.

3.4 The k^{th} component of β is β_k which is the coefficient of the k^{th} explanatory variable x_k, not including the intercept term. It represents marginal effect of a change in x_k on the average number of occurrences. If x_k is continuous,

$$\frac{\partial E(Y|\boldsymbol{X})}{\partial x_k} = \frac{\partial \lambda}{\partial x_k} = \beta_k e^{(\alpha+\boldsymbol{X}'\beta)} \quad (4)$$

$$\Rightarrow \frac{\partial \lambda/\lambda}{\partial x_k} = \beta_k.$$

If x_k is discrete and with a change in value from b to a,

$$\frac{\Delta E(Y|\boldsymbol{X})}{\Delta x_k} = E\big(Y|\boldsymbol{X}_{\backslash \boldsymbol{k}}, x_k = a\big) - E\big(Y|\boldsymbol{X}_{\backslash \boldsymbol{k}}, x_k = b\big). \quad (5)$$

3.5 A major assumption and problem of POIREG is the equidispersion (the equality of mean and variance) which is very often violated with real world data. One common remedy is to introduce so-called unobserved heterogeneity between different observations.

3.6 Negative binomial regression is just POIREG with unobserved heterogeneity introduced into the conditional mean equation (3). In brief, the mean λ_i is no longer constant but is assumed to be Gamma distributed. Details for the derivation can be found in Greene (2008) and Lord and Park (2010). We highlight here the major structure:

$$Prob(Y = y_i|\theta, \boldsymbol{X}_i) = \binom{y_i + \theta - 1}{\theta - 1}\left(\frac{\theta}{\lambda_i + \theta}\right)^{\theta}\left(\frac{\lambda_i}{\lambda_i + \theta}\right)^{y_i}, \qquad (6)$$

$$E(y_i|\theta, \boldsymbol{X}_i) = \lambda_i, \qquad (7)$$

$$Var(y_i|\theta, \boldsymbol{X}_i) = \lambda_i + \frac{\lambda_i^2}{\theta}, \qquad (8)$$

where $\theta > 0$ is a parameter of the introduced Gamma distribution which dictates the extent of the heterogeneity. λ_i, as before, is the mean of the count variable.

3.7 It is obvious from (8) that overdispersion exists as the variance exceeds the mean, a phenomenon typically found in real world data. Interpretation of the regression coefficients are as in the POIREG case.

4. The Results

4.1 The next thing is to figure out how the intervention measures can be modeled. One option is to use an indicator variable to pinpoint the announcement dates documented in paragraph 2.7. This is logical but may not capture the lasting effect of the policy between two announcement dates. We use instead an indicator variable, $\mathbb{I}_{10\cdot 2009}$, that takes on value 0 for dates before October, 2009 and a value 1 thereafter.

4.2 The indicator variable $\mathbb{I}_{10\cdot 2009}$ is then multiplied element-by-element to each explanatory variable mentioned in paragraph 2.3 and to the intercept. Pooling together a constant, the original set of explanatory variables, and the set of interactive terms just described gives the input $\alpha + \boldsymbol{X}'\beta$ required for the NB2REG. The idea is that there is *a priori* level shift and change in slopes of the regression function.

4.3 We begin the exercise with a test of choosing NB2REG over POIREG. This is equivalent to a test of $1/\theta = 0$, see equation (8). We run the likelihood ratio test:

$$LRT = -2(\mathcal{L}_R - \mathcal{L}_U) \sim \chi^2(q) \qquad (9)$$

where $\mathcal{L}_R$ and $\mathcal{L}_U$ are the log-likelihoods for the restricted model and the unrestricted model respectively, and q is the number of restrictions imposed. In our context, the null hypothesis is $1/\theta = 0$ and this means that POIREG is the restricted model while the NB2REG is the unrestricted model. The p-value of this *LRT* is 0, suggesting a rejection of the null hypothesis and an acceptance of the NB2REG model.

4.4 To see if the government intervention has a general impact on LU, we run the likelihood ratio test again with a null hypothesis of $\beta_R = 0$. β_R is the subset of the regression coefficients corresponding to the interactive terms described above. Intuitively, the restriction states that these interactive terms are insignificant and the intervention measures have nothing to do with the mortgage lending policy of banks. The p-value for the *LRT* is 0.22%, so the full model is statistically significant at both the 1% and 5% level.

4.5 Count data regression models have no built-in R^2 that offers an absolute measure of model fitness. Instead, it is a likelihood based indicator, the deviance, which is often relied on to see how well the model is doing. The *smaller* the deviance, the better the model fits. A pseudo-R^2 can be computed using ratios of deviances of the fitted model and a baseline model, and this can be interpreted pretty much the same way as the conventional R^2. Technical details of these can be

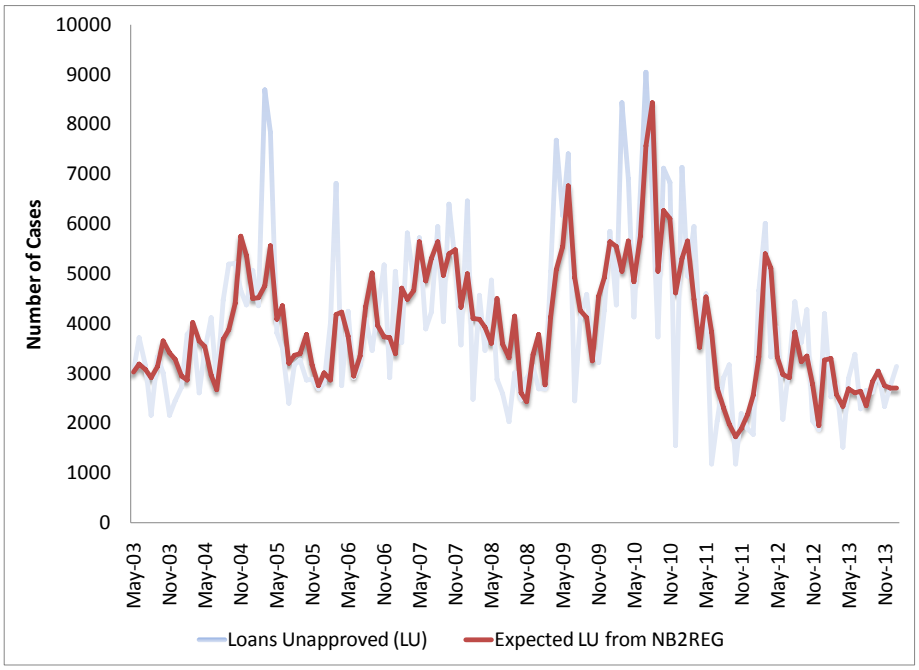
found in Cameron and Trivedi (1998). Table 1 reports the NB2REG output for the full model including estimated coefficients and the summary statistics; while Figure 6 compares the predicted mean LU with the actual data.

Table 1. Output of the Negative Binomial Regression

Variables in NB2REG	Estimated Coefficients	t-statistic
Dependent Variable:		
Loans Unapproved (LU)		
Independent Variables:		*critical value = ±1.9802*
Constant	9.6557*	4.3084
LTV Ratio (LTV)	0.0920*	2.8216
Contractual Life (CL)	-0.0306*	-4.2745
Property Price Change (ΔP)	0.0023	0.6014
S&P	0.0001*	4.2940
HIBOR-3	0.0362	1.0052
Share of Co-financing (COF)	-0.0242	-1.2324
Share of Primary transactions (PRIM)	0.0065	0.9654
Loans Written Off in 12 months (WOFF)	-0.0009*	-3.8350
Herfindahl Index (HHI)	0.0000	1.6679
$\mathbb{I}_{10\cdot2009} \times$ constant	-9.5647*	-2.8201
$\mathbb{I}_{10\cdot2009} \times$ LTV	0.0458	0.8983
$\mathbb{I}_{10\cdot2009} \times$ CL	0.0300*	3.1923
$\mathbb{I}_{10\cdot2009} \times$ ΔP	-0.0006	-0.0680
$\mathbb{I}_{10\cdot2009} \times$ S&P	-0.0000	-1.0390
$\mathbb{I}_{10\cdot2009} \times$ HIBOR-3	-0.3231	-0.3625
$\mathbb{I}_{10\cdot2009} \times$ COF	-0.1410	-0.7326
$\mathbb{I}_{10\cdot2009} \times$ PRIM	0.0049	0.3494
$\mathbb{I}_{10\cdot2009} \times$ WOFF	-0.0028	-1.1634
$\mathbb{I}_{10\cdot2009} \times$ HHI	0.0000	0.4832
Summary Statistics:		
Deviance	130.5820	
Pseudo-R^2	0.5520	

Remarks: An asterisk indicates significance at the 5% level.

Figure 6. Fit of the Negative Binomial Regression Model



4.6 In terms of model fitness, the diagram shows that the fitted means $\hat{\lambda}_i$ are less volatile than the actual data but manage to track the latter fairly well. The pseudo-R^2 indicates that about 55% of the variation in the data is explained by the NB2REG.

4.7 Recall that the parameter vector includes essentially the coefficients of the explanatory variables plus an incremental coefficient for the same variables which apply only to post October, 2009 months. We summarize here the findings and the interpretation for the explanatory variables:

- The exponential of the constant and its interactive term with the structural break dummy gives the average number of unapproved loans when all other variables equal zero. Both of these parameters are statistically significant, providing support for a level shift as a result of the government intervention. The extent of the shift is

rather enormous, from $e^{9.6557} \approx 15611$ before the break to $e^{(9.6557-9.5647)} \approx 1.095$ after the intervention.

- The coefficient of LTV is significant but that of its interactive term is not. Both of them have the same sign, meaning that the marginal effect intensifies after the break date. The pre-intervention period observes a marginal effect of $0.092 \times e^{(\hat{\alpha}+\bar{\boldsymbol{X}}'\hat{\beta})} \approx 359$ cases, and the corresponding effect post-intervention is $(0.092 + 0.0458) \times e^{(\hat{\alpha}+\bar{\boldsymbol{X}}'\hat{\beta})} \approx 492$ cases. So, for a 1% point increase in the average LTV, there will be on average 359 more rejected applications before the intervention and 492 more rejected cases after the intervention, other things being the same. Banks seem to be averse to higher leverage and are more so with the presence of government intervention.

- The coefficients of CL and its interactive term are both significant. The signs are opposite implying the marginal effect is being neutralized (reduced) after the break date. The pre and post intervention marginal effects, assuming a 1 year (12 months) increase in the average contractual life, are $E\left(Y\middle|\bar{\boldsymbol{X}}_{\setminus CL}, CL = \overline{CL} + 12\right) - E\left(Y\middle|\bar{\boldsymbol{X}}_{\setminus CL}, CL = \overline{CL}\right) \approx -1198$ cases and ≈ -23 cases, respectively. A longer contractual life encourages loan approval and reduces rejection but the impact is much smaller after the intervention.

- Interestingly, the effect of a change in property prices is insignificant to loan approval. For a 10% increase in property prices, the pre and post intervention response of LU is an additional 90 cases and 60 cases respectively.

- We have seen from Figure 1 that loans approved are highly correlated with S&P transactions. It turns out that the correlation with unapproved loans is also positive. A 1000 case increase in S&P will result in an average increase of rejection by 273 cases before intervention and 122 cases after intervention.

- Loan approval reacts differently to cost of funding before and after the intervention. Before October, 2009, each 1% drop in 3 month HIBOR reduces unapproved loans by 141 cases while the impact is an *increase* in LU by 1026 cases after the break date. Banks are therefore much less aggressive in low interest rate environment, other things the same.

- Finally, an increase in the share of approved loans with co-financing reduces the chance of rejecting loan applications before and after the break. On the other hand, an increase in the share of approved loans related to primary market transactions increases the chance of loan disapproval before and after the break.

5. Conclusion

5.1 This paper applies count data regression techniques to evaluate banks' mortgage lending behavior with and without government intervention. The amount of unapproved loans is associated with different property market and financial sector variables by way of a negative binomial regression model. Government's cooling measures, first implemented in October, 2009, are modeled as an indicator variable that interacts with other explanatory variables in our framework. The overall evidence suggests different bank lending behavior over the two periods, which may not be accounted for by the government intervention entirely. Banks are rather risk averse even in the pre intervention period, and the worldwide financial crisis may have changed the ecology of financial intermediaries as well.

5.2 Ours is a relatively simple model which did not touch up, for instance, on the issue of feedback between bank lending and property prices or transactions. More advanced count data models are available and, when employed, can enrich the scope of our analysis.

Reference

Cameron, A.C., Trivedi, P.K. (1998) *Regression Analysis of Count Data.* Econometric Society Monograph No. 53, Cambridge University Press.

Greene, W. (2008) *Functional forms for the negative binomial model for count data.* Economics Letters, 99, 585-590.

Lord, D., Park, B.J. (2010) *Appendix D – Negative binomial regression models and estimation methods.* In Crime Stat Version 3.3 Update Notes, Part 2: Regression Modeling, The National Institute of Justice, Washington DC.

Winkelmann, R. (2008) *Econometric Analysis of Count Data.* 5th Edition, Springer-Verlag.