The Park Place Economist

Volume 26 | Issue 1 Article 13

2018

Analyzing the Mortgage Loan Value and Associated Risk for a Commercial Bank

Brandon Fricke
Illinois Wesleyan University, bfricke@iwu.edu

Analyzing the Mortgage Loan Value and Associated Risk for a Commercial Bank

Abstract

In this paper I will analyze the loan amount to income level of individuals applying for mortgage loans in the years following the financial recession. Additionally I will examine the impact that each independent variable has on the overall loan amount. Underwriting looks at many of these variables when analyzing a loan to determine a borrowers credit history. Here I will attempt to quantify the impact of some of these variables. I will use time series analysis techniques to examine any trends and patterns within the ratio. It is hypothesized that as applicant income increases, the loan amount applied for will increase at an even larger percentage, as applicants now have more peace of mind and financial security.

Analyzing the Mortgage Loan Value and Associated Risk for a Commercial Bank

Brandon Fricke

I. Introduction

It may seem hard to believe, but the modern mortgage as we know it today has only been around since 1934. Under the new Federal Housing Administration (FHA) guidelines, they set up programs to offer individuals loans for 80 % loan-to-value and higher (How Mortgages Work 2002). Loan-to-value, as defined by Investopedia(2017), is the amount of the mortgage lien divided by the appraised value of the property. This created more opportunities for the average American to own a home, and forced commercial banks to adopt similar lending policies. Initially this allowed the typical American, who couldn't afford to make a down payment of 70 to 80 % under the old banking guidelines, to now qualify for a mortgage loan. However this new practice made mortgage lending far more risky; homeowners had far less invested into their homes, and were taking out increasingly larger loans that many could not afford. Financial institutions were searching for another measure to identify credit worthy customers and their ability to repay the mortgage loan. The debt-to-income ratio, or more simply the debt ratio, became a popular way to quantify risk, and measures an individual's debt payment to overall income.

Data obtained from the Bureau of Labor Statistics and CoreLogic indicated that debt to income ratios were increasing in the years leading up to the housing market collapse of 2007. Banks and lenders alike were attempting to qualify more Americans for home loans that they could not afford all in an effort to boost their bottom line. Mortgage debt figures obtained from the Federal Reserve board of St. Louis showed that mortgage debt was rapidly rising in the years leading up to the housing bubble burst. With an increase in debt, and thus increased levels of risk, defaults and foreclosures could prove catastrophic to the housing market, which is exactly what happened. Homeowners with mortgage loans that they could not afford because they took on too much risk began to default and banks began to repossess their houses. The bubble burst and the housing market began to bottom out with the number of individual defaults in the millions per year (Equifax, 2017).Tighter regulations on mortgage loan qualifications were subsequently put in place, including stricter regulations on debt to income levels of individuals looking for a mortgage loan. A number of factors, including shock from the recent collapse and strict regulations set out by the United States government and other Government-Sponsored Enterprises forced debt to income ratios of individuals to decrease in the following years. Mortgage lenders

seemingly began to issue loans only to those individuals deemed credit worthy and demonstrated the ability to repay their home loan.

Along with credit score, the debt ratio is one of the primary factors in deciding if an individual will qualify for a loan. It is a crucial measure of overall loan risk that financial institutions look for (CFPB 2017). If the average loan value to income level is increasing it could signal individuals and financial institutions taking on more risk in financing mortgages. Evidence from the Consumer Finance Protection Bureau suggests that borrowers with a higher debt to income ratio are more likely to run into trouble making monthly mortgage payments. As recent history has shown us, high risk loans in the housing market can prove disastrous to the worldwide economy. As stated above, it is known that the loan amount to income ratio was increasing nationwide up until 2007 when it began to decline. In this paper I will analyze the loan amount to income level of individuals applying for mortgage loans in the years following the financial recession. Additionally I will examine the impact that each independent variable has on the overall loan amount. Underwriting looks at many of these variables when analyzing a loan to determine a borrowers credit history. Here I will attempt to quantify the impact of some of these variables. I will use time series analysis techniques to examine any trends and patterns within the ratio. It is hypothesized that as applicant income increases, the loan amount applied for will increase at an even larger percentage, as applicants now have more peace of mind and financial security.

II. Data and Methods

In order to test my hypothesis relating applicant income to loan amount, data was obtained from a commercial bank in Topeka, Kansas. This unique time series analysis study will focus on the bank's mortgage lending practices from 2010 until 2017. Data on applicant's income and the corresponding loan amount will encompass the greater Topeka, Kansas geographical area. The loan amount and applicant income figures I will analyze have no specific frequency as loan volumes have fluctuated throughout the past seven years. From January of 2010 to September of 2017, the range of data I will be analyzing, the volume of loans per month and even per day fluctuated due to numerous factors. With a total of approximately 3000 loans to analyze over the past seven years we have an average of 400 loans each year. Additionally loan amount and applicants income are recorded in thousands of dollars, thus making it simpler to examine.

As previously stated, there is no specific frequency to the data analyzed as each day and month have different loan volumes. Mortgage lending is a heavily seasonal industry according to the National Association of Realtors, and this trend is displayed in my data as well. During the winter months, typically November to January, loan volume is considerably

lower than during the late spring and summer months. Because of this discrepancy it will be important to test for seasonality to statistically confirm our assumptions as well as seasonally adjust the series so we can accurately analyze trends and patterns. Graph One takes all 3000 data points and averages each month's loan amounts and corresponding income values resulting in a total of 93 months to evaluate. When looking at the graph there is a small but noticeable upward trend in both variables values over the past seven years. This could be due to a number of factors including easing of credit requirements and economic recovery as I move farther and farther away from the financial crisis and housing meltdown of 2007-2008. Another point to note is that as the applicants stated income fluctuates, the loan amount applied for closely resembles a similar pattern with a slight lag associated with it. However, both series show a considerable amount of volatility in them, making it difficult to predict future values.

Because of the volatility and seasonality associated with my data I experience a wide range of maximum and minimum values. Average loan amount, excluding loan and income figures above 500, fluctuates from 38 to 238 with applicant's income showing similar fluctuations. One possible explanation to this is a few data outliers that are present in our two series, which distorts the mean for those months. When looking at the raw unadjusted data there are four loan amounts above 1000, or $1,000,000 since I am viewing the variables in thousands of dollars. By removing these outliers I can better analyze the data and view any patterns present, excluding the statistical anomalies. Graph Two shows a scatter plot of loan amount to total income with a logarithmic regression line fitted to the series. A large majority of the data points are loan amounts below 500 and stated incomes below 500. Graph Three now shows a scatter plot of the data points with the loan amounts and stated income levels above 500 excluded. As one can see, the fitted logarithmic regression line shows a similar trend to that of diminishing marginal returns. As income increases loan amount increases as well, but only to a certain extent as shown by the regression line flattening out. This is in contradiction to my previous hypothesis, that loan amount would increases exponentially relative to increased income.

For this study I will be using the following estimation equation to analyze any correlation between loan amount and income level:

$$\Delta \text{ Loan Amount} = \alpha + \beta\Delta \text{ Income} + \beta\Delta \text{ Time}$$

The change in loan amount is dependent upon a constant factor plus the change in the independent variables applicant income and time. With an improving economy and a recovering housing market income as well as loan amount would be expected to rise. Additionally, as income increases consumers will have greater purchasing power and the increased ability

to take a larger mortgage loan. Because the economy continues to improve, the housing market has been recovering and annual inflation has been increasing. The variable time is expected to have a positive correlation with loan amount. Thus the dependent variable and both independent variables are expected to take positive values. Limitations for this analysis come from the volatility of the series as well as the seasonality component which will be corrected when I adjust the variables for seasonality.

In this study I will use EViews software to analyze the data and use ordinary least squares to run a regression. Additionally I will use the ADF and KPSS tests to check for unit roots and stationarity in the series. Testing for unit roots and stationarity, a common time series practice, will determine the order of integration of our series. Finally I will test the residuals to ensure they behave normally and would not adversely affect my regression results.

III. Literature Review and Theory

The correlation between loan amount and stated income values has been heavily analyzed since the housing crisis of 2008 because it is a strong measure of risk associated with the mortgage loan. The risk from the bank's perspective has been analyzed with the leverage ratio, or capital to assets that a bank uses to fund a loan. Furlong and Keeley studied the relationship between bank's leverage ratio and the risk associated with their portfolio investments (Furlong & Keeley 1989). They found that as capital increased, and thus the leverage ratio also increased, banks began to invest in more prudent assets which reduced overall default risk. A higher leverage ratio is thus considered safer for banks as they must use more of their own capital to finance loans or sell off the most risky assets and investments.

A 2010 study by Adrian and Shin found evidence supporting Furlong and Keeley's original hypothesis. Their study furthered the original work and found evidence of pro-cyclical patterns in the capital to asset ratio. The pro cyclical nature of the leverage ratio acts as an amplification mechanism to business cycles upturns and downturns (Adrian & Shin 2010). As asset values increase from an improving economy a bank's leverage ratio falls and they will have a surplus of capital. Banks will then purchase additional assets or securities to more than just restore the previous leverage ratio. An upward pressure on asset prices follows and thus the upward spiral begins. Equally, a negative shock to asset prices would trigger a downward spiral of leverage. Leading up to the 2008 financial crisis asset prices and mortgage-backed securities were rising rapidly, thus lowering banks' leverage ratios. Banks began to make riskier investments into the sub-prime market granting credit to borrowers who did not have the ability to repay. What followed was the subsequent financial crisis and downturn in the business cycle.

These findings are in contradiction to the results of previous studies from Koehn that found the opposite effect; an increase in the leverage ratio was associated with an increase in investment risk (Koehn, 1979). His explanation was that as the capital-to-asset ratio rises, banks can afford to give out fewer loans as their capital is tied up elsewhere. Thus they would be incentivized to invest in more risky assets to maintain a profit level. More recent studies from Lee and Hsieh found similar results, and concluded that the parallel relationship between capital and risk levels was due to moral hazard and bank managers not doing enough to avoid unnecessary risk (Lee & Hsieh, 2013).

Furthermore, numerous studies have been conducted in determining if a loan will be approved and how much debt can be issued to a borrower. Loan to value, according to Investopedia, is one of the primary measures if a loan will be approved. In 1992 Vandell used a bi-proportional hazard model to examine risk associated with mortgage loans (Vandell, 1992). He found that at the individual borrower level, loan-to-value was the primary determinant of mortgage default. Banks denying loans based on inadequate loan to value ratios are highly correlated with a reduction in mortgage defaults. In 1994 Abraham furthered the original findings and concluded that while loan-to-value is highly correlated with default rates, the debt-coverage ratio has greater explanatory power in determining mortgage defaults. Similar to Vandell's study, my study examines the risk associated with individual borrowers at the microeconomic level. Additionally I am also analyzing the loan to income ratio of individuals to examine default risk, which has proven to be a good measure.

Underwriting standards are the final measure in approving or denying a loan. It will determine how much debt can be issued, the terms of the loan, and analyzes the borrowing history of customers. Underwriters must determine if a borrower can repay the loan, analyzing income, employment, and bank statements. Based on these factors as well as the credit history of the borrower will determine if the loan amount applied for will be approved. In 1996, Brent Ambrose developed a model of the market for commercial real estate loans based on these variables used in property decision making (Ambrose 1996). Underwriting standards, according to Ambrose, determine the supply price of mortgage loans. He also found that as lines of credit tighten, and thus underwriting standards tighten, the loan to value ratio is forced to decrease. With a larger down payment by consumers banks can now finance the loan using less of their own capital and still maintain adequate leverage ratios, thus ensuring they make prudent investments.

Many of the above studies analyze macroeconomic data with a wide scope and scale. My study is more focused to microeconomic data analyzing a single commercial bank. Similar to Ambrose's 1996

study, I am attempting to quantify how much these variables that underwriting examines affect the overall loan amount. Adding to the above theories I will analyze the risk associated with the banks mortgage lending strategies and attempt to calculate what variables impact the loan amount. I will determine if the commercial bank analyzed is increasing their risk levels by granting loans to less credit worthy borrowers, indicated by an increasing spread in loan amount to applicant income.

IV. Results

As shown in Graph One, loan amount and applicant income have proven to be volatile series since 2010. Some of this volatility could be due to the heavy seasonality component associated with mortgage lending. As stated before, mortgage lending is a heavily seasonal industry according to the National Association of Realtors. To compensate for this, I have seasonally adjusted both the loan amount and applicant income using the Census X12 method and the seasonal means for both variables can be seen in Graph Four. The final regression equation will examine seasonally adjusted variables.

Before running the OLS regression we must ensure that the series is stationary and does not contain a unit root. To do so I will run standard time series tests, namely the Augmented Dickey Fuller (ADF) and Kwiatkowski–Phillips–Schmidt–Shin (KPSS) tests to ensure my series is stationary and would not produce spurious results. Table One shows the ADF and KPSS tests for the variable applicant income. I can, at the 99% confidence level, reject the null hypothesis of the ADF test and assume that the series does not contain a unit root in levels. Furthermore when examining the KPSS test I cannot confidently reject, and thus must accept, the null hypothesis that the series is stationary in levels. Table Two again runs the same ADF and KPSS tests now on the variable loan amount and finds similar results. I can again assume, with 99% confidence, that the series does not contain a unit root in levels and that the series is stationary. Additionally, because I can confidently assume there is no unit root present and the series is stationary in levels I say my series is integrated of order zero, I(0).

With all of the variables stationary and the preliminary transformations completed it is now possible to run a regression using ordinary least squares. The basic regression equation generated is as follows:

$$LOAN = 24.566 + 1.058\ INCOME + 0.2838\ TIME + 112.6\ DUMMY$$

Seasonally adjusted loan amount is the dependent variable. It is a function of a constant plus applicant's income as well as time and a dummy variable. Again, the variables are in terms of thousands and the coefficients are thus reported accordingly. Additionally I have added a dummy variable to the regression

equation for December of 2016. Graph Five shows the residual diagnostics when no dummy variable was implemented and it is clear that the residual for this time period takes a value well outside the standard deviation. Thus I have implemented a dummy variable for December of 2016 to remove the residual data outlier.

Table Three outlines the regression results in detail with the total observations, or months analyzed, being 93. The constant term of the equation indicates that when all other variables take a value of zero, the loan amount will still be $24,566. Applicant's income, as expected, had the greatest impact on loan amount out of all the variables analyzed. For every $1,000 increase in applicant's income, the loan amount would increase by $1,058. More simply, a $1.00 increase in income results in a $1.06 increase in loan amount. Furthermore, the income variable is significant at the 99% confidence level. Additionally I have added a variable to capture how the loan amount will change over time. Time, measured in months, is statistically significant at the 99% confidence level. As each month passes the loan amount will increase by $284. This could be due to a number of factors including an improving economy and inflation increasing. Time, while not as impactful as income, still has a significant influence on the loan amount. The adjusted R-squared measures the variability in the dependent variable as explained by the independent variables in the model, but adjusts the statistic based on the number of independent variables used. For this regression the adjusted R-squared took a value of 0.5266, meaning that 52.6% of the variability in loan amount is explained by our independent variables. The F-statistic, which assumes that all coefficients actually have a value of zero, can be confidently rejected and thus we can assume our variables have explanatory power.

We must also test the residuals of the regression to ensure they have a constant variance, show no signs of autocorrelation, and behave normally. When measuring for autocorrelation we use the Breusch Godfrey test which has a null hypothesis that the residuals are not auto-correlated. Because the p-value falls outside our range of confidence we cannot reject the null hypothesis, and thus must assume that the residuals do not show signs of autocorrelation. Additionally we must use the Jarque Berra test to examine the normality of the residuals. The null hypothesis for this test is that the residuals are normally distributed. Because the P-value falls well outside our range of confidence we cannot reject the null hypothesis and thus must assume that the residuals behave normally. Finally we must test for heteroscedasticity, or that the residuals have a constant variance, using White's test. Because our p-value is above 0.05 we can assume that our residuals are homoscedastic. Overall our residuals show no signs of autocorrelation, are normally distributed, and have constant variance. Thus our residuals should not adversely affect our regression results.

V. Conclusions

As one can see from Graph One, applicants income closely follows the pattern and trend of loan amount. This broadly indicates that the bank is not resorting to risky mortgage lending practices, as the ratio of applicant's income to loan amount is relatively stable. If the spread between loan amount and applicant income was increasing it would have signified borrowers taking on more debt and increasing risk to both parties. Additionally loan amount does not increase exponentially as income increases. In fact this study shows that as income is increasing loan amount also increases, but only to a certain extent. These findings could be due in part by bank size and lending limits.

As stated before, the data collected for this study was obtained from a commercial bank in Topeka, Kansas. Because there were varying amounts of loans issued each month no two months were the same. For simplicity I averaged each month's loan amount and income to have a total of 93 data points from January 2010 to September of 2017. Due to the seasonality and low volume of loans in the winter months I removed all seasonal components of the data using the Census X12 method. Thus the regression equation used included the seasonally adjusted variables. Before running a regression I tested the data for unit roots and stationarity using the ADF and KPSS tests respectively. Both variables, loan amount and applicant income, showed no unit root and were stationary in levels. As such the series is said to be integrated of order zero, I(0).

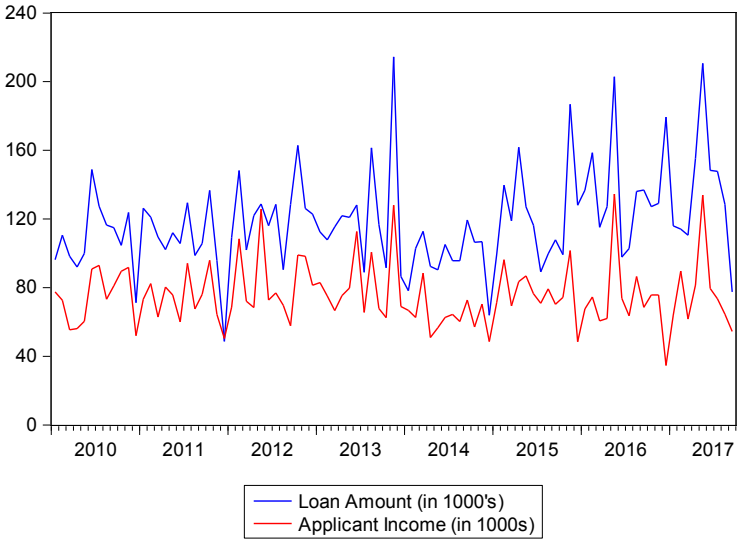
As anticipated, the variable of applicant income had the largest effect on overall loan amount and was statistically significant. For every $1,000 increase in income the loan amount will rise by $1058. Loan amount is also a factor of time, measured in months, and was statistically significant. As each month passes the loan amount will increase by $284 holding all other variables constant. This could be due to inflation, the improving economy, a rising local housing market or a combination of the three. Additionally the constant term indicated that when all other variables took a value of zero the loan amount would still be $24,566. The dummy variable implemented was again used to remove the residual outlier from December 2016. I can confidently reject the F-statistic as well, and assume that the coefficients take a value other than zero. Finally, the adjusted R-square took a value of 0.5266 indicating that the independent variables explain 52.6% of the variance in the loan amount. The residual diagnostics showed positive results, indicating no autocorrelation and normally distributed residuals with a constant variance.

In 1992 Vandall studied the relationship between the loan-to-value and default risk in mortgage lending. He found that an increase in the ratio was strongly correlated with mortgage default. Addition-

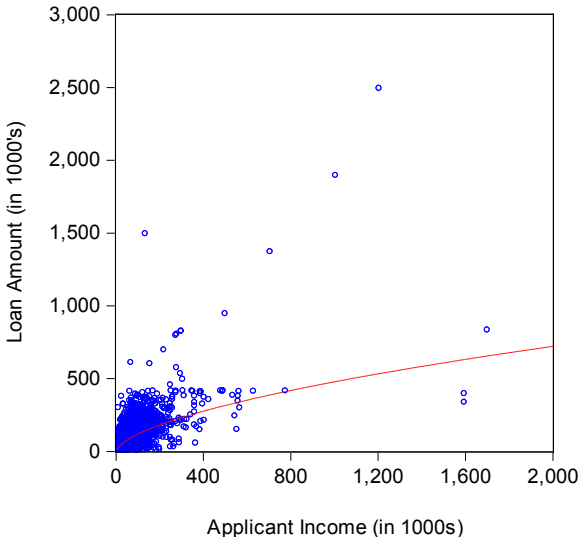
ally, Ambrose found that underwriting standards will help determine the amount of debt issued. Here I am examining risk associated with mortgage lending, both to the bank and the consumer. While underwriting standards are unique to each lender and personalized to each borrower, I am attempting to quantify the loan amount using two independent variables. The findings could be useful in numerous ways. Bank and lending underwriters could use this to see how much of an impact these variables have on overall loan amounts. To further this study a number of changes and additions could be made. To gain further explanatory power on the loan amount additional variables could be added to the equation. For instance, credit score could be included as an independent variable in attempting to explain the loan amount. Additionally bank size could be a contributing factor to the loan amount, and may be a reason why we see a glass ceiling when applicant income is increasing. Further analysis of the local housing market, both pre-recession and post-recession, could provide more insight into lending conditions and help explain the variables trends and patterns. These variables could be implemented in the future to gain further knowledge of what effects the loan amount while analyzing the local housing market to see its affects on lending.

Appendix

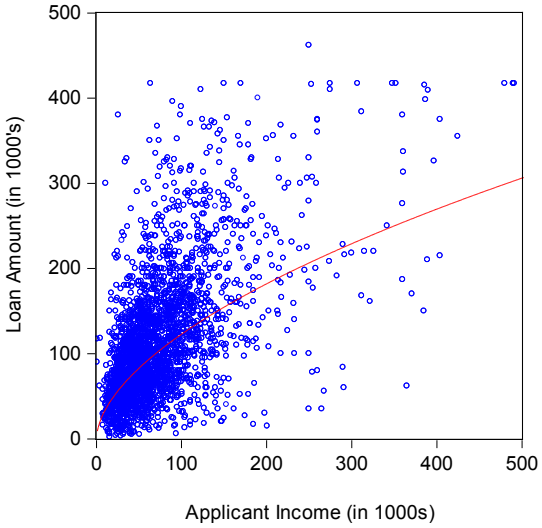
Graph One



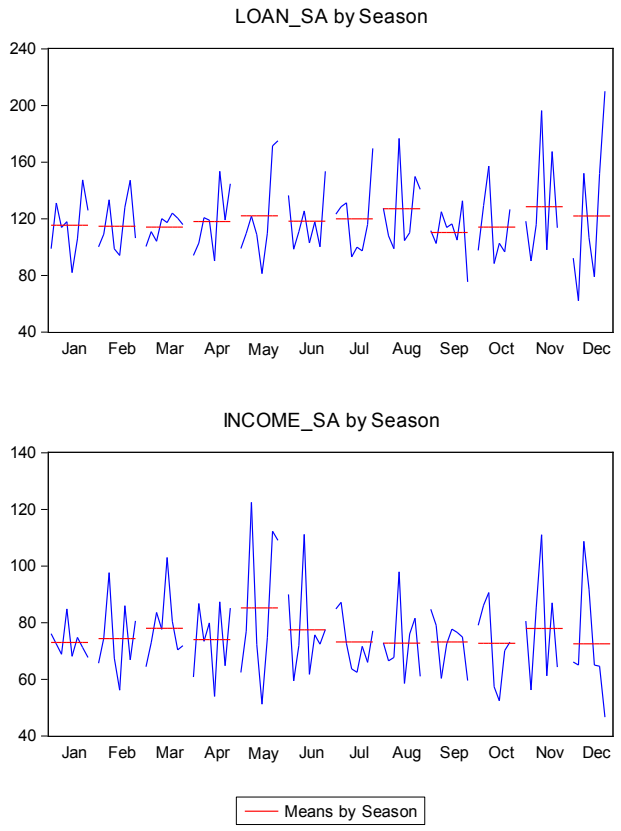
Graph Two



Graph Three



Graph Four



Graph Five:

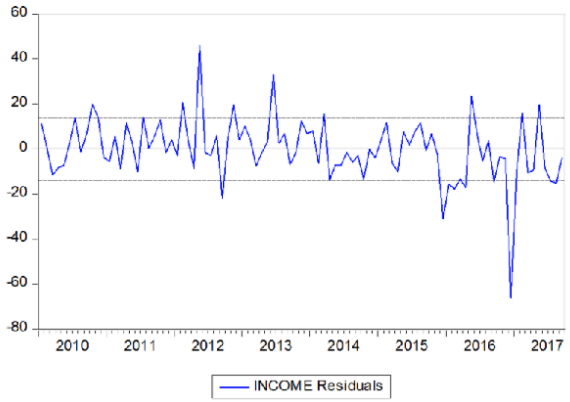


Table One.

Applicant's Stated Income

		Levels: Constant, Linear Trend	f.o.d.:Constant
Augmented Dickey-Fuller test statistic		T-Statistic: -9.7743	T-Statistic: -14.1871
Test Critical Values	1%	-4.0608	-3.5047
	5%	-3.4593	-2.8939
	10%	-3.1557	-2.5841

Applicant's Stated Income

		Levels: Constant, Linear Trend	f.o.d. Constant
Kwiatkowski-Phillips-Schmidt-Shin test statistic		t-statistic: 0.0584	t-statistic: 0.0261
Test Critical Values	1%	0.216	0.739
	5%	0.146	0.463
	10%	0.119	0.347

Table Two.

Loan Amount

		Levels: Constant, Linear Trend	f.o.d.:Constant
Augmented Dickey-Fuller test statistic		T-Statistic: -8.9535	T-Statistic: -7.6391
Test Critical Values	1%	-4.0608	-3.5083
	5%	-3.4593	-2.8955
	10%	-3.1557	-2.5849

Loan Amount

		Levels: Constant, Linear Trend	f.o.d. Constant
Kwiatkowski-Phillips-Schmidt-Shin test statistic		t-statistic: 0.1065	t-statistic: 0.0148
Test Critical Values	1%	0.216	0.739
	5%	0.146	0.463
	10%	0.119	0.347

Table Three.

Tabulation of Regression Results

Dependent Variable: Seasonally Adjusted Loan Amount, N=93

Constant	24.566**
	2.294
Income	1.058***
	8.062
Time	0.2838***
	4.041
DUMMY	112.61***
	6.038
Adj. R-Squared	0.5266
s.e. of equasion	17.988
F-Statistic	35.113***
Residual Diagnostic Test:	
Autocorrelation	1.5587
Normality of residuals	0.332
Heteroskedasticity	2.2042*

References:

Ambrose, B., & Benjamin, J. (spring 1996). Credit Restrictions and the Market for Commercial Real Estate Loans. *Real Estate Economics, 24*(1). Retrieved November 15, 2017.

Beccalli, E., Boitani, A., & Giuliantonio, S. (2015). Leverage pro-cyclicality and securitization in US banking. *Journal of Financial Intermediation, 24*, 200-230.

Brain, S. (2017, May 10). Home Foreclosure Statistics 2016. Retrieved September 13, 2017, from http://www.statisticbrain.com/home-foreclosure-statistics/

C. (2017, March 3). What is a debt-to-income ratio? Why is the 43% debt-to-income ratio important? Retrieved November 13, 2017, from https://www.consumerfinance.gov/ask-cfpb/

what-is-a-debt-to-income-ratio-why-is-the-43-debt-to-income-ratio-important-en-1791

C. (2013, January 10). Consumer Financial Protection Bureau Issues Rule to Protect Consumers from Irresponsible Mortgage Lending. Retrieved November 13, 2017, from https://www.consumerfinance.gov/about-us/newsroom/consumer-financial-protection-bureau-issues-rule-to-protect-consumers-from-irresponsible-mortgage-lending/

Federal Reserve Economic Data | FRED | St. Louis Fed. (n.d.). Retrieved November 13, 2017, from https://fred.stlouisfed.org/

Folger, J. (2017, March 14). Debt-To-Income Ratio - DTI. Retrieved September 13, 2017, from http://www.investopedia.com/terms/d/dti.asp?optly_redirect=integrated&lgl=myfinance-dynamic-tags

Furlong, F., & Keeley, M. (1989). A Reexamination of mean variance analysis of bank capital regulation. *Journal of Banking and Finance, 14*(1), 69-84.

Furlong, F., & Keeley, M. (n.d.). Capital Regulation and Bank Risk-Taking. *Federal Reserve Bank of San Francisco, 13*, 34-39.

Investopedia. (2017, February 15). Loan-To-Value Ratio - LTV Ratio. Retrieved September 13, 2017, from http://www.investopedia.com/terms/l/loantovalue.asp?optly_redirect=integrated&lgl=myfinance-dynamic-tags

Koehn, M., & Santomero, A. (december 1980). Regulation of Bank Capital and Portfolio Risk. *The Journal of Finance, 35*(5), 1235-1244.

N. (n.d.). Www.nar.realtor. Retrieved November 13, 2017, from https://www.nar.realtor/

Obringer, L. A., & Roos, D. (2002, October 08). How Mortgages Work. Retrieved September 13, 2017, from http://home.howstuffworks.com/real-estate/buying-home/mortgage.htm

Oyedele, A. (2017, July 08). America's biggest mortgage source is making it easier for millennials to buy their first home. Retrieved November 13, 2017, from http://www.businessinsider.com/millennials-homeownership-fannie-mae-debt-to-income-ratio-mortgages-2017-7

Repullo, R. (january 2006). Capital Requirements, Market Power, and Risk-Taking in Banking. *Journal of Financial Intermediation, 15*(1), 160.

Rosa, R. (2016, April 8). Massachusetts Home Buyer Guide. Retrieved November 13, 2017, from http://info.buyersbrokersonly.com/blog/what-is-debt-to-income-ratio-for-massachusetts-homebuyers

Ross, S. (2014, November 14). How do leverage ratios help to regulate how much banks lend or invest? Retrieved November 13, 2017, from https://www.investopedia.com/ask/answers/111414/how-do-leverage-ratios-help-regulate-how-much-banks-lend-or-invest.asp

Vandell, K.D. 1992. Predicting Commercial Mortgage Foreclosure Experience. Journal of the American Real Estate and Urban Economics Association 20: 55-88