Local Real Estate Markets

Leverage and House Price elasticities

Yichuan Wang[1]

University of Michigan Undergraduate Research Opportunities Program

March 12, 2013

[1] Faculty Sponsor: David Bieri

Outline

Previous Literature

- Basic Hypothesis: Areas with higher proportions of high loan to value (LTV) households have higher house price volatility
- Lamont and Stein (1999) – proportion of houses with loan to value (LTV) ratio above 80% (high LTV) is strong predictor of response of house prices to income shocks
  - Theoretical result from a model of collateral constraints to explain price-volume correlation
- Baseline specification:

$$\Delta P_t = \beta_0 + \underbrace{\beta_1 \Delta Y_t + \beta_2 \Delta d_t}_{\text{controls}} + \underbrace{\beta_3 \Delta d_t \times \Delta P_t}_{\text{effect of LTV}} + \underbrace{\beta_4 \Delta P_{t-1}}_{\text{momentum}} + \underbrace{\beta_5 \frac{P_t}{Y_t}}_{\text{mean reversion}}$$

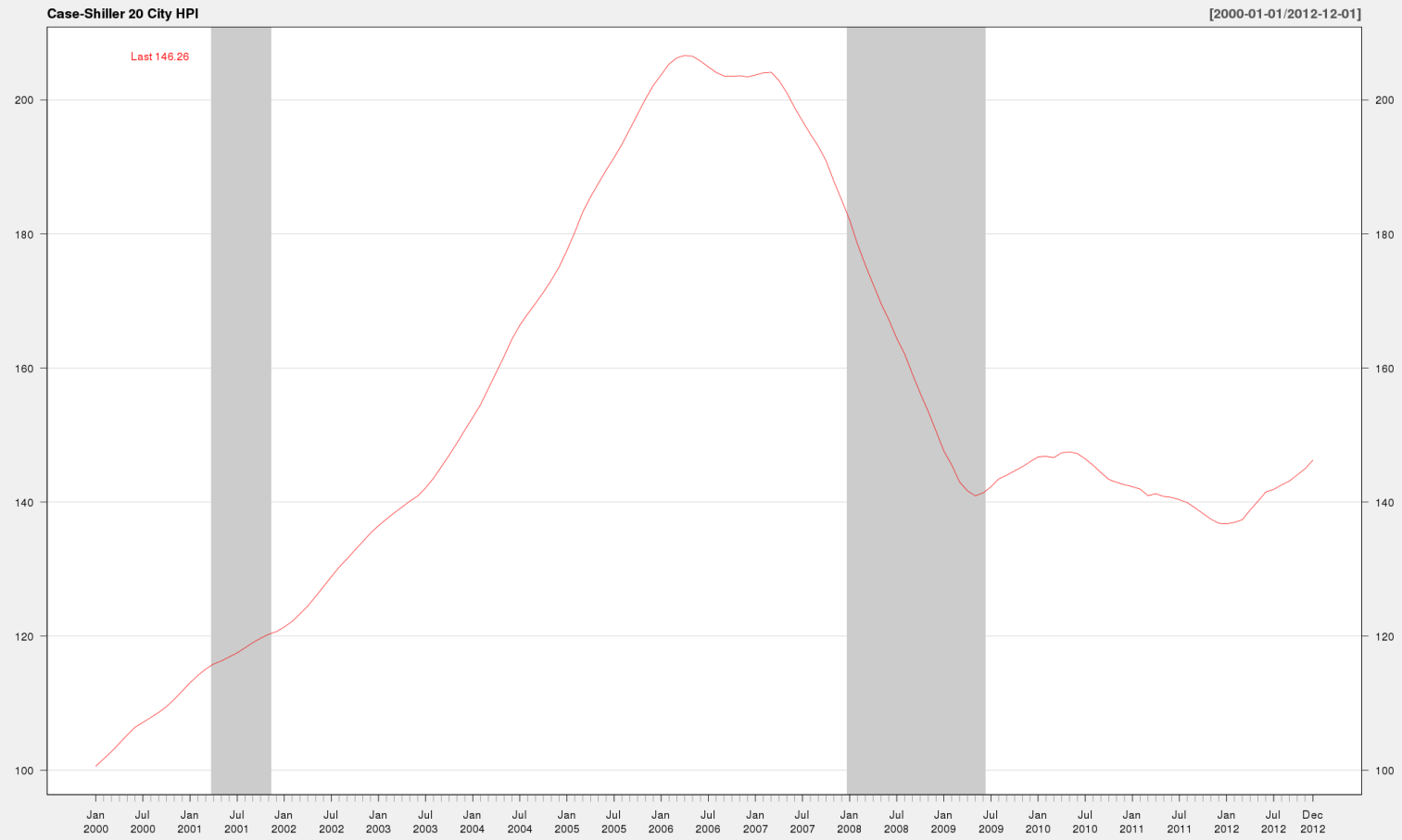
P	House Price
Y	Per Capita Income
d	Proportion of high LTV

New Studies and Papers

1. New Literature:
   - Mian and Sufi (2009) – Subprime mortgages and mortgage defaults
   - Saiz (2008) – New techniques to estimate land and house price increases
2. New Data – American Housing Survey, extremely large 2011 dataset
   - Also includes pure (un-estimated) values of LTV

Housing Bubble

- Very severe drop in home prices
- Goal is to try to explain what happened



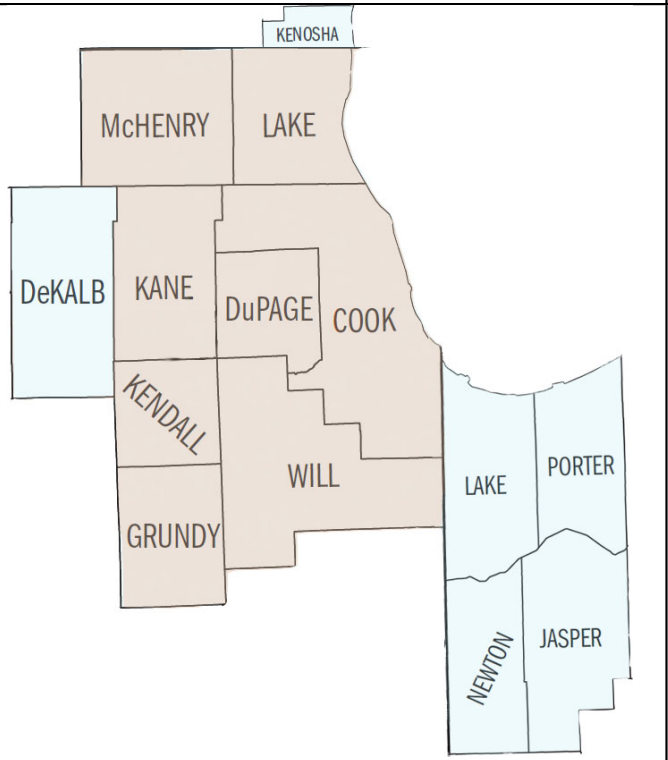
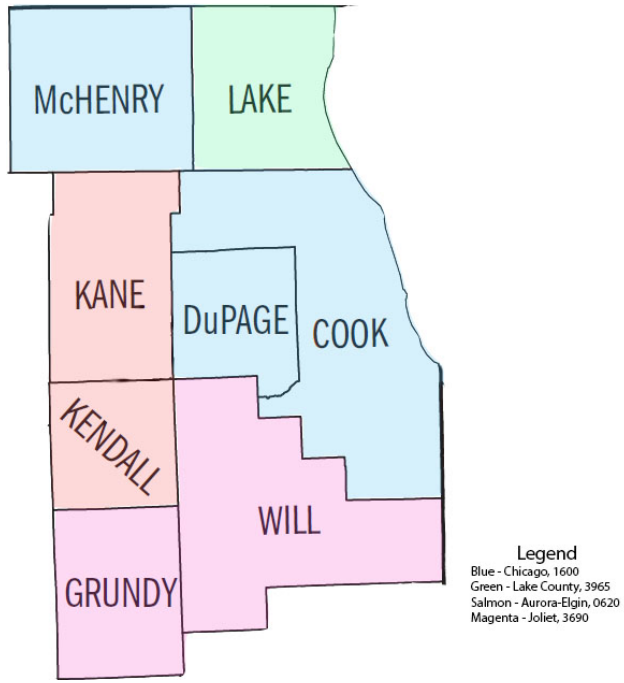
Outline

Data Sources

- Loan to Value - calculated from the American Housing Survey (AHS)
  - Metropolitan Survey that takes place roughly every 2 years
  - Survey households, both "owner occupied" and rental
  - Look at data from 1995–Present
- House Prices - Freddie Mac House Price Index
- Per Capita Income, Population – BEA Local Personal Income (CA1-3)

Geographical Divisions

- What does it mean to be in the Chicago Metropolitan Area?
- Consider:

Chicago-Joliet-Napiersville	Chicago AHS
	

- Process all the data with Java, Python

Outline

Difficulties with Previous Approach

- Previously, HUD would use complex estimation – very hard to replicate
- Get errors that are several orders of magnitude greater than expected:
- Summary statistics for difference between R implementation of HUD code and the actual survey value:

	clpeva	Rclpeva
Min	0.00	-2604
1Q	44.00	0
Median	72.00	0
Mean	77.61	3611
3Q	95.00	54
Max	5034.00	23639695
NA's	144940	131699

- One redeeming value is that the number of NA's are similar

Estimation Procedure

- Simplify the problem
  - We follow Lamont and Stein and set all households with LTV $> 80\%$ as "highly leveraged households"
  - Let $L = 1$ for a highly leveraged household, and $L = 0$ otherwise
  - Now it becomes a binary choice – much less information under consideration
- Logit Regression:

$$\Pr(L|X) = \frac{1}{1 + e^{-\beta X}}$$

β	vector of regression coefficients $(\beta_0, \beta_1 \dots)$
X	vector of explanatory variables $(1, x_1, x_2 \dots)$

- Declare as "highLTV" if model predicts over 0.5

Variables Selected

- Final variables considered:

Variable	Description
jmmort	Whether the first mortgage value was edited by HUD - topcoding or hotdeck
pmiamt	Amount of mortgage insurance paid
AdjValue	Value of Property in dollars divided by 100,000[2]
mcnt	Number of Mortgages
AdjAmmort	Amount of the first mortgage in dollars divided by 100,000

- Potential variables that were excluded: Rclpeva, canvar, jvalu

[2] Scaling to get meaningful regression coefficients

Outline

Estimation Results

	Dependent variable:
	highLTV
jmmort	-0.392^{***}
	(0.126)
pmiamt	0.882^{***}
	(0.123)
AdjValue	-2.037^{***}
	(0.081)
mcnt	1.360^{***}
	(0.151)
AdjAmmort	2.631^{***}
	(0.101)
Constant	-1.841^{***}
	(0.181)
Observations	4,009
Log likelihood	$-1,787.962$
Akaike Inf. Crit.	3,587.925
Note:	*p<0.1; **p<0.05; ***p<0.01

Contingency Table

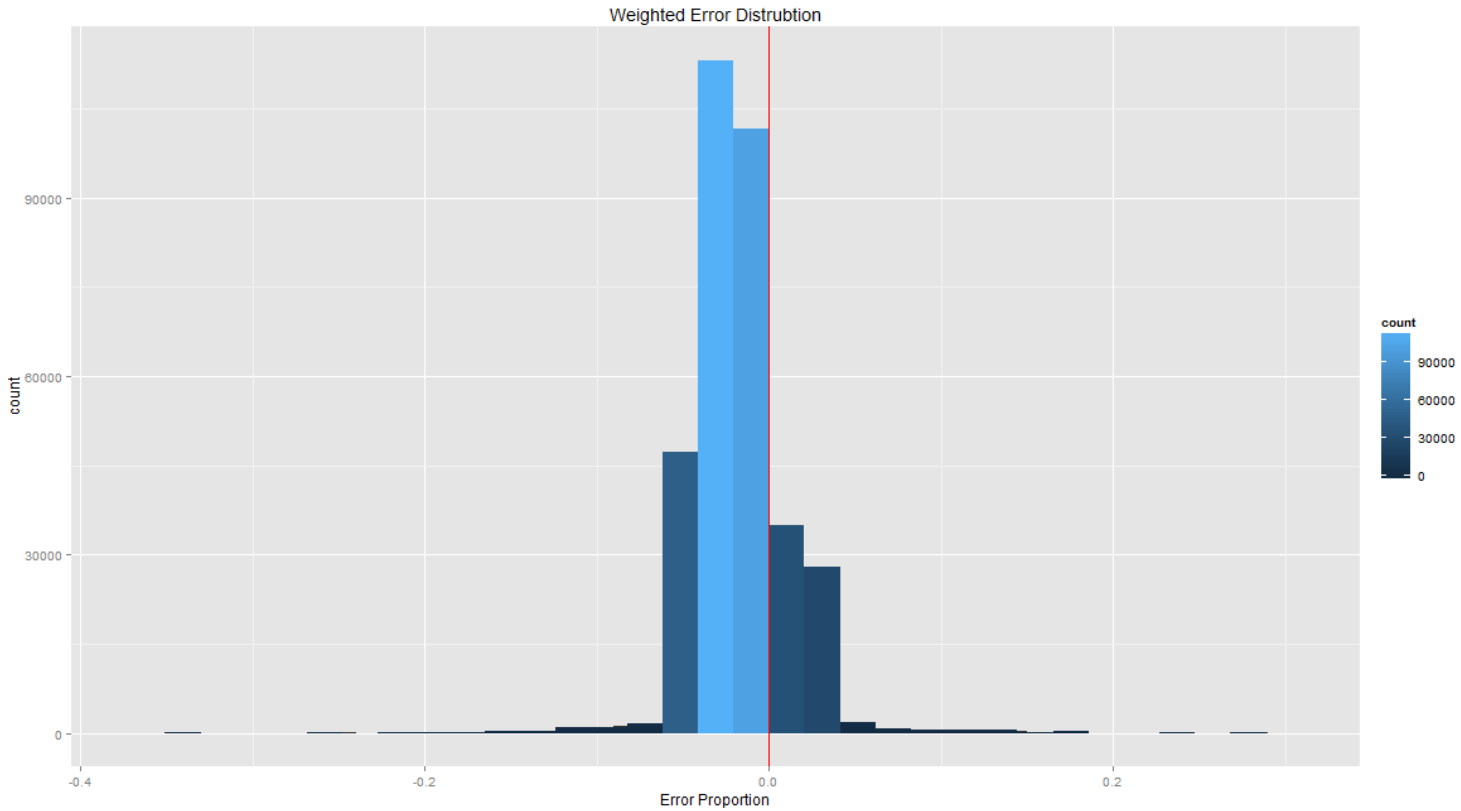
- Run 10 trials – train the regression on 10% of the data, then test performance on the entire dataset
- Use small training set to test 'out of model' consistency

		High LTV (AHS)		
		No	Yes	Row Totals
High LTV (Model)	No	213686	34069	247755
	Yes	27154	140171	167325
	Column Total	240840	174240	

- Sensitivity: 80%, Specificity: 88%

Error Distribution

- Similar rate of Type I, Type II errors suggests that model can be quite accurate – only need total number of LTV in an SMSA
- For each SMSA, error can be much smaller – after weighting by the count of households (red line is 0)



New Tools

- New Tools
  1. More accurate table of counties of the AHS
  2. Unified database linking BEA per capita income data to the AHS
  3. Regression model to predict which households are highly leveraged
- Implications
  - Whether a household is highly leveraged is *consistent* and reasonably *predictable*
  - Can make later housing finance research using the AHS much simpler and more robust

Outline

Bibliography

Lamont, O. and J. C. Stein (1999). Leverage and house-price dynamics in U.S. cities. *RAND Journal of Economics*.

Mian, A. and A. Sufi (2009). The Consequences of Mortgage Credit Expansion: Evidence from the U.S. Mortgage Default Crisis. *The Quarterly Journal of Economics*.

Saiz, A. (2008). On local housing supply elasticity. *Wharton Working Paper*.