

Subprime crisis influence on house ownership rate by races

By James Chiu
Jiqi Pan
Wenyi Cai

1. Introduction

- Back ground information: subprime crisis in 2007 to 2009.

Proportion of Subprime Home Loans by Race (2006)

African American families (Black)	Hispanic families (Latino)	white non-Hispanic families (White)
52.44%	40.66%	22.20%

Source: [Center for Responsible Lending](#) analysis of 2006 Home Mortgage Disclosure Act (HMDA) data reported by the [Federal Financial Institutions Examination Council](#)

- Incentive of the paper: to study the subprime crisis's influence to different races.

2.Literature review

- Race and home Ownership,1900-1990
By *William J. Collins and Robert A. Margo* (NBER working paper)
- Mortgage Pricing Differentials Across Hispanic, Black, and White Households: Evidence from the American Housing Survey (February 2006) (HUD USER paper)

Race and home ownership, 1900-1990

- The paper use home ownership rate data of black and white from 1900 to 1990.
- The paper use linear probability model to estimate the variables' influence to the home ownership rate. The variables including race, gender, age, geography area, and whether own a land.
- The study further do decomposition study to explain the black and white home ownership gap, in some period the change of the gap is result from the political reasons.
- This paper is our project's example, but the timeline we use is different from the paper

Mortgage Pricing Differentials Across Hispanic, Black, and White Households: Evidence from the American Housing Survey (February 2006)

- Published by Office of Policy Development and Research (PD&R) of HUD center
- Use metropolitan area samples of the American Housing Survey (AHS) for 1998, 2002, and 2004 data
- focus to examine that the differences in the interest rates obtained by homeowners of different races and income levels can be explained by differences in characteristics of the borrowers, the property, and the loan itself.

3. Data

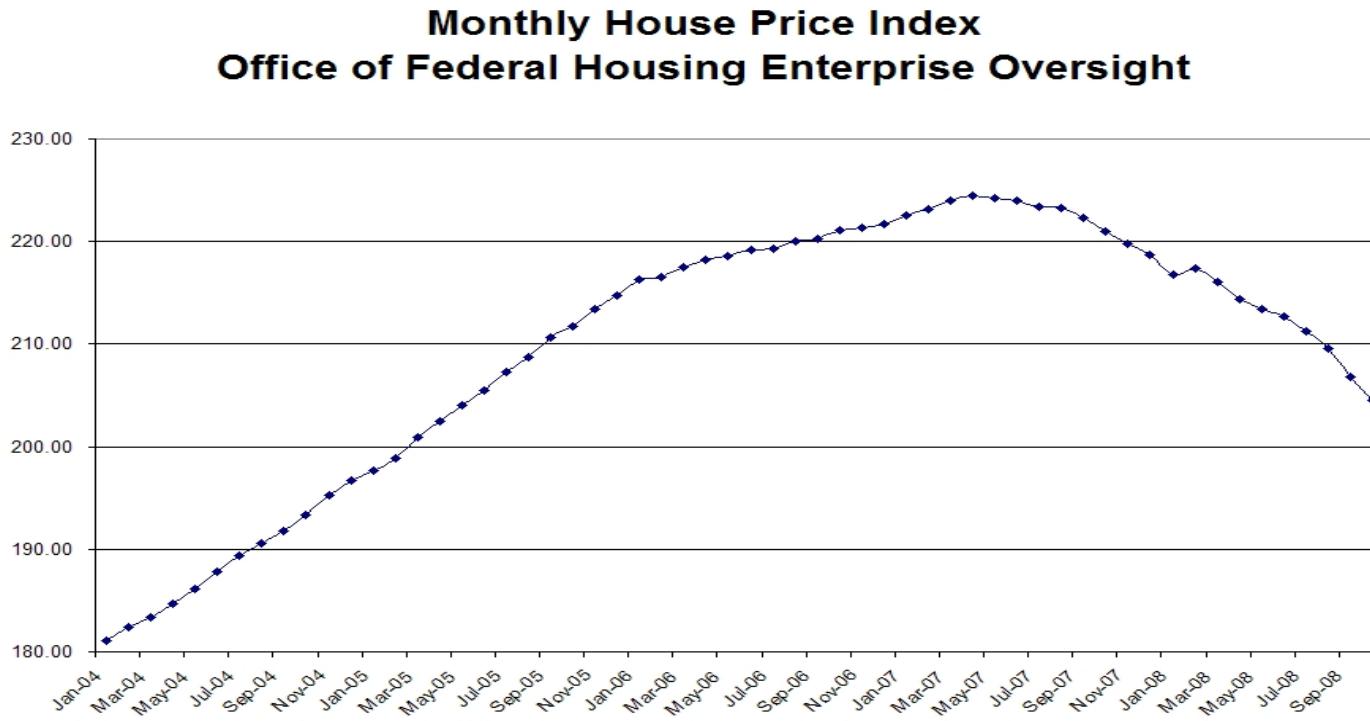
- Data source: Kauffman index dataset 2007 and 2008 The Kauffman Index measures the monthly rate of business creation at the individual owner level.
- Data size: 678703 observations in each year's data set.

3. Data

- California is the third biggest state in the United State, and has a large number of minority population. Our study only include the observations who residents in California.

4. Data dividing (housing price)

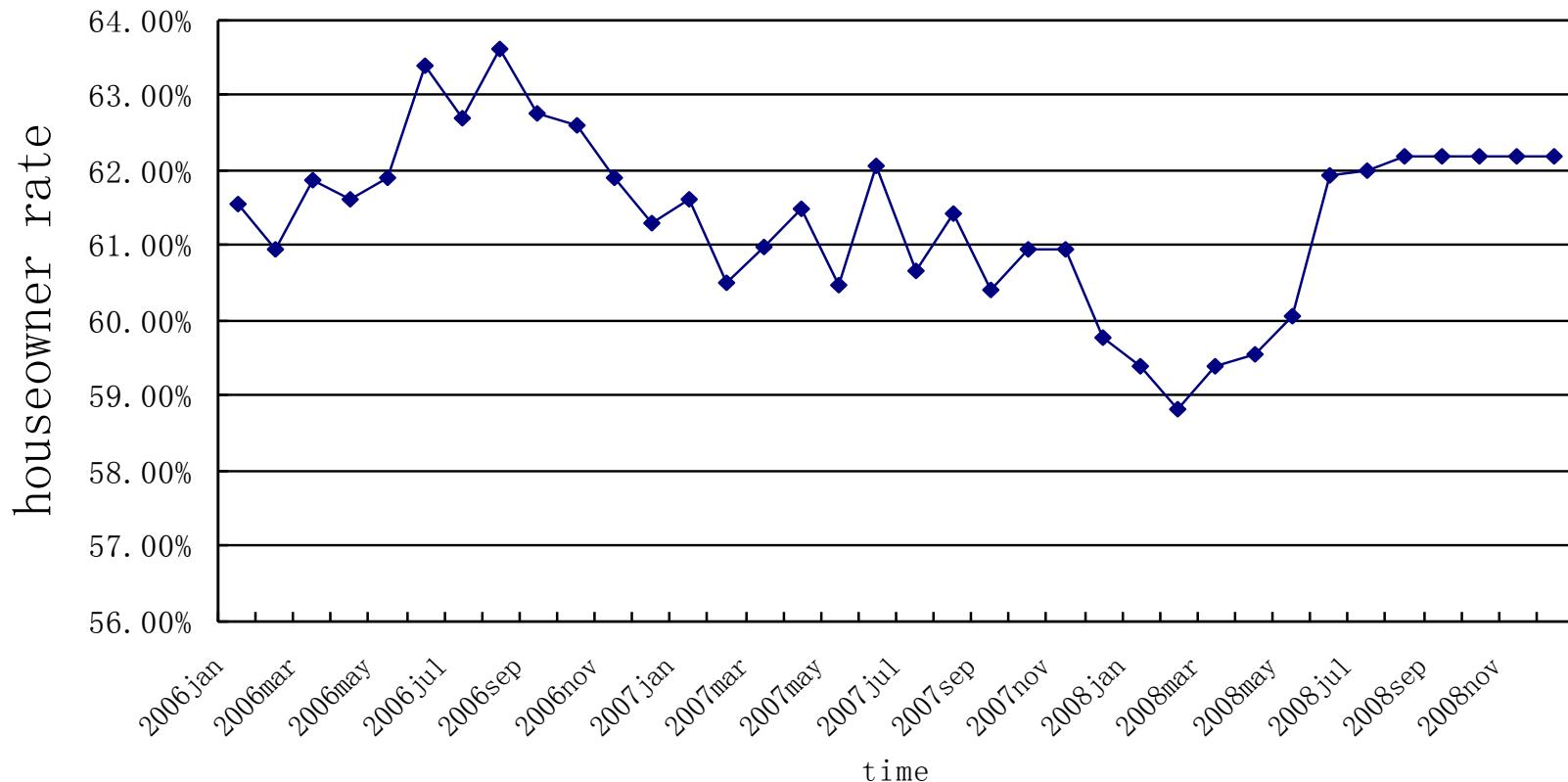
- The house price line from Jan2004 to Sep2008.



- the housing price reached to its highest point in the May 2007 and it started fast dropping in September 2007, We select May 2007 to divide the pre-subprime period and post-subprime period.

4. Data dividing (house owner rate)

the homeownership rate change from 2006 to 2008 in California



There is a 7 months lag for the house ownership rate to react to the housing price fall.

5. Variables

Dependent variable:

Homeown

Homeown=1 if own a house

Independent Variables:

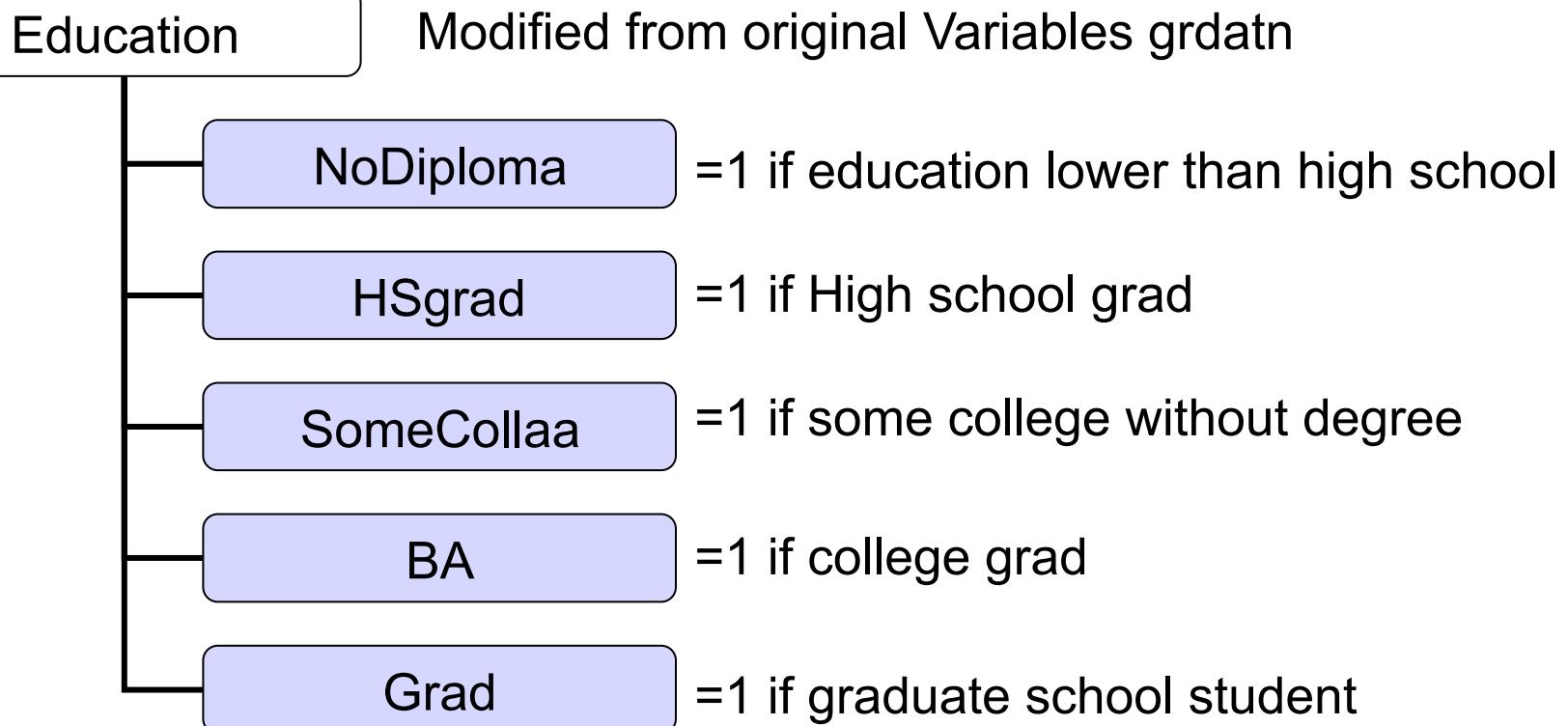
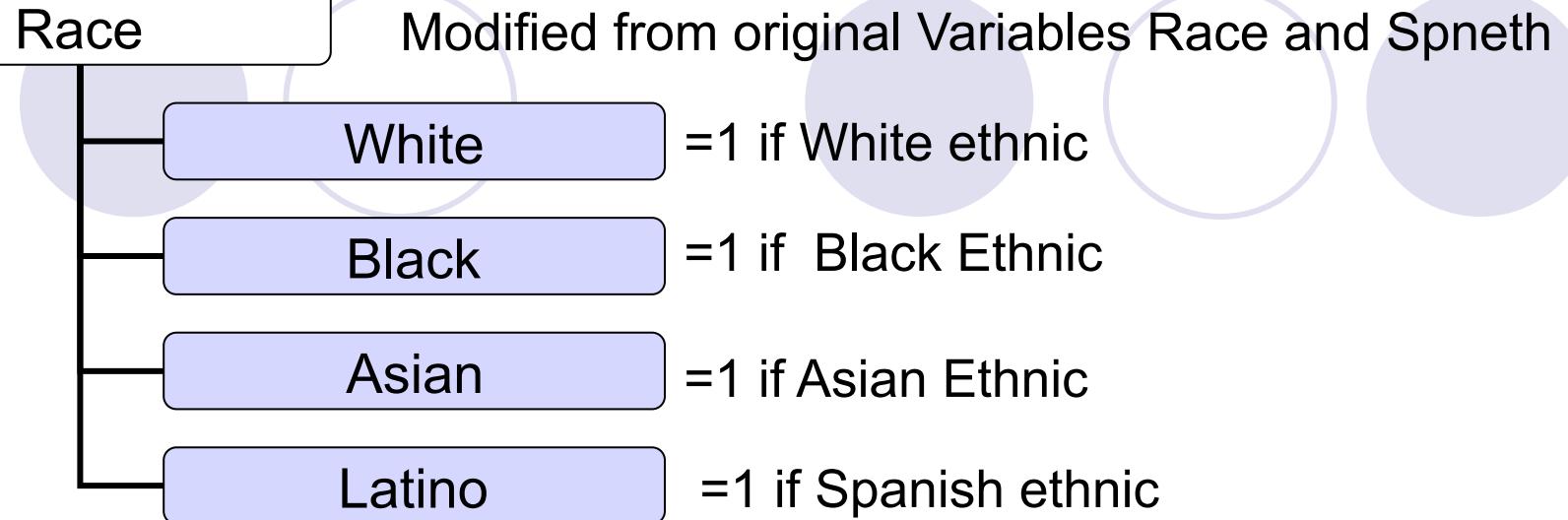
Age

Gender

Male=1 Female=0

Martial status

Married=1 if currently married



Income

Modified from original variable faminc

Incomelessthan30000

=1 if income less than 30000

Incomelessthan60000

=1 if income less than 60000 but
higher than 30000

Incomelessthan100000

=1 if income less than 100000
but higher than 60000

Income100000higher

=1 if income higher than 100000

Employment status

Modified from original variable Mrl_t1

employed

=1 if employed in the monthly data

unemployed

=1 if unemployed in the monthly data

NILF

=1 if disable retire or others

Class of Worker

Modified from original variable Class_t1

GovWorker

=1 if employ by government

Private

=1 if employ by private organization

Selfemploy

=1 if self employed

Central City Status

Modified from original variable Central City Statue

Centralcity

=1 if live in the central city area

BalanceMSA

=1 if live in suburbs

Nonmetropolitan

=1 if live far from metropolitan area

Noidentified

=1 if observation not identified

Mean of variables in the pre-subprime period(Jan2007 to Jul2007)

Variable	N	Mean	S.D.	Variable	N	Mean	S.D.
homeown	64950	0.61813	0.485	Incomelessthan60000	64950	0.2330	0.4222
age	64950	40.8525	12.11	Incomelessthan100000	64950	0.4249	0.4942
male	64950	0.48217	0.499	Income100000higher	64950	0	0
Married	64950	0.58823	0.492	Employed	64950	0.7365	0.4396
White	64950	0.78965	0.408	Unemployed	64950	0.0341	0.1801
Black	64950	0.05421	0.226	NILF	64950	0.2292	0.4197
Asian	64950	0.12110	0.326	GovWorker	64950	0.10799	0.3110
Latino	64950	0.33639	0.472	Private	64950	0.5553	0.4968
NoDiploma	64950	0.17248	0.378	SelfEmploy	64950	0.1051	0.3068
HSGrad	64950	0.22809	0.419	centralcity	64950	0.16694	0.3442
SomeCollAA	64950	0.30494	0.460	BalanceMSA	64950	0.15496	0.3335
BA	64950	0.20146	0.401	Nonmetropolitan	64950	0.00666	0.0738
Grad	64950	0.09301	0.246	notidentified	64950	0.48217	0.1571
Incomelessthan30000	64950	0.19199	0.393				

Mean of variables in the post-subprime period (Jan2007 to Jul2007)

Variable	N	Mean	S.D.	Variable	N	Mean	S.D.
homeown	60782	0.60402	0.48906	Incomelessthan60000	60782	0.22403	0.41694
age	60782	41.0916	12.2810	Incomelessthan100000	60782	0.42988	0.49506
male	60782	0.48494	0.49977	Income100000higher	60782	0	0
Married	60782	0.58296	0.49307	Employed	60782	0.73771	0.43987
White	60782	0.78380	0.41165	Unemployed	60782	0.04014	0.19629
Black	60782	0.05888	0.23540	NILF	60782	0.22213	0.41568
Asian	60782	0.12388	0.32945	GovWorker	60782	0.11029	0.31326
Latino	60782	0.34251	0.47455	Private	60782	0.56465	0.49580
NoDiploma	60782	0.16830	0.37414	SelfEmploy	60782	0.09999	0.29999
HSGrad	60782	0.22911	0.42026	Nopay	60782	0.00108	0.03293
SomeCollAA	60782	0.29362	0.45542	centralcity	60782	0.46146	0.49851
BA	60782	0.21037	0.40757	BalanceMSA	60782	0.43417	0.49565
Grad	60782	0.09858	0.25606	Nonmetropolitan	60782	0.01742	0.13084
Incomelessthan30000	60782	0.17457	0.3796	notidentified	60782	0.08693	0.28174

Logit Model

$$H_t = \beta_0 + \beta_1 X_t + \beta_2 I_t + \beta_3 G_t + \varepsilon_t$$

- H: homeownership rate (Binary dependent variable)
- X: Ethnicity dummy variables
- I: a vector of individual characteristics
- G: Geographical dummies
- β : coefficient vectors
- ε : error term
- t: periods (pre May 2007 and post May 2007)

Ethnicity Categories

- Whites*
- Blacks
- Asians
- Latinos
- * indicates that this group is the reference group in the regression

Individual characteristics

- age
- working hours
- family income
 - 0-30,000 *
 - 30,000-60,000
 - 60,000-100,000
- education
 - No high school Diploma*
 - High school graduate
 - Some college
 - BA
 - Graduate (MA&Doctorate)
- employment status
 - Employed*
 - Unemployed
 - Not in labor force
- class of worker
 - Government employee*
 - Private employee
 - Self-employed
- Gender
- marital status

Geographic Location

- Central City*
- Metropolitan statistical area, but not in a central city
- Non-metropolitan area
- Not identified

Full Scale Regression

Dependent Variable : Homeownership Rate Dummy

Variables	Pre-May 2007		Post-May 2007	
	Coefficient	P-value	Coefficient	P-value
Intercept	-1.4675	<.0001	-1.9517	<.0001..
Ethnicity..				
Black	-0.4037	<.0001	-0.5390	<.0001..
Asian	0.1152	0.0003	0.1679	<.0001..
Latino	0.0184	0.5347	-0.0553	0.0052..
Individual Characteristics..				
age	0.0370	<.0001	0.0402	<.0001..
Single	-0.6245	<.0001	-0.5884	<.0001
male	-0.0348	0.0565	-0.0217	0.2833..
working hours	-0.00156	0.0356	-0.000246	0.0014..
Income..				
Incomelessthan60000	0.3648	<.0001	0.3425	<.0001..
Incomelessthan100000	1.4589	<.0001	1.4956	<.0001..
Education..				
HSGrad	0.4945	<.0001	0.4324	<.0001..
SomeCollAA	0.6968	<.0001	0.6777	<.0001..
BA	0.5484	<.0001	0.5188	<.0001..
Grad	0.5637	<.0001	0.4242	<.0001..
Employment Status..				
Unemployed	-0.0649	0.2785	-0.0950	0.3237..
NILF	-0.3521	<.0001	-0.3171	<.0001..
Class of Worker..				
Private	-0.4764	<.0001	-0.4063	<.0001..
SelfEmploy	0.0598	0.1386	-0.0317	0.0092..
Geographic Location..				
BalanceMSA	0.3500	<.0001	0.6699	<.0001..
Nonmetropolitan	0.7547	<.0001	0.8678	<.0001..
Notidentified	0.3810	<.0001	0.4856	<.0001..
Observations	64950		60782..	

Education's Effect on homeownership rate by races

Pre

	white	black	Asian	Latino	
Pre	Hsgrad	0.4912	0.9782	0.6242	0.4298
	somecollege	0.7339	1.0416	0.9129	0.6225
	BA	0.5168	1.2567	0.7795	0.6078
	Grad	0.6230	1.3741	0.6364	0.766
Post	Hsgrad	0.5142	0.5023	0.6447	0.3268
	somecollege	0.7939	0.5559	1.041	0.6226
	BA	0.5280	0.7545	0.8114	0.2972
	Grad	0.5273	1.0345	0.4119	0.64

Post

Homeownership Regression With Different Control Variables

White is the reference group

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Coefficients:	Pre	Post										
Black	-0.4037 (0.0393)	-0.5390 (0.0407)	-0.5405 (0.0385)	-0.6904 (0.0389)	-0.3965 (0.0397)	-0.5269 (0.0404)	-0.4793 (0.0394)	-0.6095 (0.0401)	-0.3728 (0.0393)	-0.4937 (0.0399)	-0.4211 (0.0392)	-0.6197 (0.0402)
Asian	0.1152 (0.0291)	0.1679 (0.0307)	-0.0181 (0.0286)	-0.02 (0.0294)	0.1014 (0.0295)	0.1612 (0.0306)	0.1710 (0.0293)	0.2405 (0.0304)	0.0358 (0.0291)	0.0715 (0.03)	0.0954 (0.0290)	0.0714 (0.0304)
Latino	0.0184 (0.0223)	-0.0553 (0.0241)	-0.1571 (0.0221)	-0.2198 (0.0230)	-0.1862 (0.0208)	-0.2414 (0.0216)	0.0961 (0.0228)	0.0123 (0.0238)	-0.1898 (0.0223)	-0.2574 (0.0232)	0.0106 (0.022)	-0.0772 (0.0238)
Control Variables:												
Income	Yes		No		Yes		Yes		Yes		Yes	
Education	Yes		Yes		No		Yes		Yes		Yes	
Marital Status	Yes		Yes		Yes		No		Yes		Yes	
Age	Yes		Yes		Yes		Yes		No		Yes	
Geographic Location	Yes		No									
Others	Yes											

Standard errors are in parentheses.

The Change of Homeownership Rate During the Crisis

	pre	post	change
race	mean	mean	
White	0.6188	0.6093	-1.53%
Black	0.5058	0.45	-11%
Asian	0.6657	0.6518	-2%
Latino	0.5186	0.501	-3%

Preliminary Conclusion

- All the four ethnic groups deteriorated during the crisis. Blacks got much worse than the other three groups. Asians still have the biggest homeownership rate. Whites are the most stable one.
- Income, education, age, marital status and MSA status are more correlated with the racial difference in the change of homeownership rate than other factors.

Decomposition Model

- **An Extension of the Blinder - Oaxaca Decomposition Technique to Logit and Probit Models - Robert W. Fairlie (Jan 2006)**
- The Blinder-Oaxaca decomposition technique is widely used to identify and quantify the separate contributions of group differences in measurable characteristics, such as education, experience, marital status, and geographical differences to racial and gender gaps in outcomes. The technique cannot be used directly, however, if the outcome is binary and the coefficients are from a logit or probit model.
- **Linear Decomposition**

$$\bar{Y}^W - \bar{Y}^B = \left[(\bar{X}^W - \bar{X}^B) \hat{\beta}^W \right] + \left[\bar{X}^B (\hat{\beta}^W - \hat{\beta}^B) \right]$$

- **Non-Linear Decomposition for Logit Model**

$$\bar{Y}^B - \bar{Y}^W = \left[\sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^W)}{N^B} - \sum_{i=1}^{N^W} \frac{F(X_i^W \hat{\beta}^W)}{N^W} \right] + \left[\sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^B)}{N^B} - \sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^W)}{N^B} \right]$$

Variables

$$\bar{Y}^B - \bar{Y}^W = \left[\sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^W)}{N^B} - \sum_{i=1}^{N^W} \frac{F(X_i^W \hat{\beta}^W)}{N^W} \right] + \left[\sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^B)}{N^B} - \sum_{i=1}^{N^B} \frac{F(X_i^B \hat{\beta}^W)}{N^B} \right]$$

- **X** – includes the variables:
 - (age, male, single, hours, faminc, educ, class, mlr and msa)
 - faminc: incomelessthan60000 incomelessthan100000
 - educ: hsgrad somecollaa ba grad
 - class: private selfemploy
 - mlr (Monthly Labor Force): unemployed nilf
 - msa (Metropolitan Status): balancems nonmetropolitan notidentified
- **β** – Coefficient Estimates
- **N** – Number of Observations

Blacks to White Decomposition

	Pre	Post	%Δ
Number of Obs	54809	51220	
N of obs G=0	3521	3579	
N of obs G=1	51288	47641	
Pr(Y!=0 G=0)	0.50582221	0.45012573	0.110
Pr(Y!=0 G=1)	0.61876072	0.60930711	0.015
Difference	-0.11293851	-0.15918138	-0.409
Total Explained	-0.03658716	-0.0598288	-0.635

Blacks to White Decomposition

homeown (pre)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	0.0057008	-0.0504770	0.0007486	7.62	0	0.0042335	0.007168
male	0.0001327	-0.0011750	0.0008714	0.15	0.879	-0.0015752	0.001841
single	-0.0404168	0.3578655	0.0034413	-11.7	0	-0.0471616	-0.033672
hours	0.0095268	-0.0843539	0.0026913	3.54	0	0.0042519	0.014802
faminc	-0.0331215	0.2932702	0.0021863	-15.2	0	-0.0374065	-0.028837
educ	0.0230958	-0.2044989	0.0042390	5.45	0	0.0147876	0.031404
class	0.0053401	-0.0472833	0.0024251	2.2	0.028	0.000587	0.010093
mlr	-0.0120331	0.1065456	0.0026453	-4.55	0	-0.0172178	-0.006848
msa	0.0058944	-0.0521912	0.0024152	2.44	0.015	0.0011606	0.010628
Total Explained	-0.0365872	0.3239565					
homeown (post)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	0.0014526	-0.009125439	0.000531	2.74	0.006	0.0004121	0.002493
male	0.0006347	-0.003987275	0.000829	0.77	0.444	-0.0009891	0.002259
single	-0.0336737	0.211542958	0.003843	-8.76	0	-0.0412053	-0.026142
hours	0.001681	-0.01056028	0.001452	1.16	0.247	-0.0011641	0.004526
faminc	-0.0375086	0.235634344	0.002588	-14.5	0	-0.0425814	-0.032436
educ	0.006731	-0.042285096	0.003448	1.95	0.051	-0.0000268	0.013489
class	0.0038626	-0.024265401	0.002493	1.55	0.121	-0.0010239	0.008749
mlr	-0.0042508	0.026704128	0.001705	-2.49	0.013	-0.0075929	-0.000909
msa	0.0013848	-0.00869951	0.003606	0.38	0.701	-0.0056835	0.008453
Total Explained	-0.0598288	0.375853005					

Asians to White Decomposition

	Pre	Post	
Number of Obs	59154	55171	
N of obs G=0	7866	7530	
N of obs G=1	51288	47641	%Δ
Pr(Y!=0 G=0)	0.66564963	0.65179283	0.021
Pr(Y!=0 G=1)	0.61876072	0.60930711	0.015
Difference	0.04688891	0.04248572	0.094
Total Explained	0.02955566	0.02011762	0.319

Asians to White Decomposition

homeown (pre)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	0.0047857	0.102064646	0.0005785	8.27	0	0.0036517	0.0059196
male	0.001081	0.023054492	0.0004064	2.66	0.008	0.0002845	0.0018775
single	0.0039666	0.084595697	0.0006324	6.27	0	0.0027271	0.0052061
hours	-0.0001506	-0.003211847	0.0003592	-0.42	0.675	-0.0008546	0.0005534
faminc	0.0061895	0.132003495	0.0008541	7.25	0	0.0045155	0.0078634
educ	0.0174585	0.37233751	0.0031422	5.56	0	0.0112998	0.0236171
class	0.0009376	0.0199962	0.000518	1.81	0.07	-0.0000776	0.0019529
mlr	-0.0018098	-0.038597613	0.0007053	-2.57	0.01	-0.0031921	-0.0004275
msa	-0.003018	-0.0643649	0.0007683	-3.93	0	-0.0045238	-0.0015122
Total Explained	0.02955566	0.630333697					
homeown (post)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	0.0069024	0.162464	0.0005948	11.61	0	0.0057367	0.0080681
male	0.0003995	0.009403159	0.0003226	1.24	0.216	-0.0002328	0.0010319
single	0.0045396	0.106850019	0.0008513	5.33	0	0.0028712	0.0062081
hours	-0.0000773	-0.001819435	0.0002739	-0.28	0.778	-0.0006142	0.0004595
faminc	0.0063254	0.148882966	0.0006526	9.69	0	0.0050462	0.0076045
educ	0.0172514	0.406051727	0.0033417	5.16	0	0.0107017	0.0238011
class	0.000293	0.006896435	0.0003837	0.76	0.445	-0.000459	0.0010451
mlr	-0.0018126	-0.042663747	0.0005317	-3.41	0.001	-0.0028548	-0.0007704
msa	-0.0136632	-0.321595115	0.0040362	-3.39	0.001	-0.021574	-0.0057524
Total Explained	0.02011762	0.473514866					

Latinos to White Decomposition

	Pre	Post	
Number of Obs	52276	48664	
N of obs G=0	988	1023	
N of obs G=1	51288	47641	%Δ
Pr(Y!=0 G=0)	0.55769231	0.48484848	0.131
Pr(Y!=0 G=1)	0.61876072	0.60930711	0.015
Difference	-0.06106842	-0.12445862	-1.038
Total Explained	-0.09080038	-0.07528803	0.171

Latinos to White Decomposition

homeown (pre)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	-0.0282417	0.462459975	0.003426	-8.24	0	-0.0349573	-0.0215262
male	0.0022985	-0.037638111	0.000942	2.44	0.015	0.0004532	0.0041438
single	-0.0014593	0.023896148	0.000768	-1.9	0.058	-0.0029652	0.0000466
hours	0.0027617	-0.045223047	0.00226	1.22	0.222	-0.0016671	0.0071905
faminc	-0.0486192	0.79614308	0.00447	-10.9	0	-0.0573807	-0.0398577
educ	-0.0177518	0.290687069	0.004963	-3.58	0	-0.0274794	-0.0080243
class	0.0015119	-0.024757477	0.002349	0.64	0.52	-0.0030923	0.0061162
mlr	-0.0017504	0.028662932	0.001883	-0.93	0.353	-0.0054412	0.0019405
msa	0.0000784	-0.001283806	0.002334	0.03	0.973	-0.0044963	0.0046532
Total Explained	-0.09080038	1.486863096					
homeown (post)	Coefficients	% of Diff Explained	Std. Err.	z	P> z	[95% Conf.]	Interval]
age	-0.0374586	0.300972323	0.005471	-6.85	0	-0.0481816	-0.0267356
male	0.0003059	-0.002457845	0.000573	0.53	0.593	-0.0008163	0.0014281
single	-0.0082584	0.066354584	0.00212	-3.89	0	-0.0124141	-0.0041027
hours	0.0019355	-0.015551354	0.001592	1.22	0.224	-0.0011847	0.0050556
faminc	-0.03184	0.255828001	0.004611	-6.9	0	-0.0408777	-0.0228023
educ	-0.0069908	0.056169673	0.006017	-1.16	0.245	-0.0187836	0.0048021
class	0.0010578	-0.00849921	0.002359	0.45	0.654	-0.0035647	0.0056803
mlr	3.28E-06	-2.63541E-05	0.002884	0	0.999	-0.0056495	0.005656
msa	0.0070818	-0.05690084	0.002944	2.41	0.016	0.0013112	0.0128525
Total Explained	-0.07528803	0.604924191					

Conclusions

- Blacks and Latino's homeownership decreased dramatically compared to whites having a (41% and 103%) increase in the gap respectively.
- While Asians actually close the gap by 9.4%.
- Blacks explained characteristics are more focused in the *mlr*, *faminc*, *single* and *educ* variables.
- The Asian gap is mostly influenced by age, *single*, *faminc*, *educ* and also *msastat*.
- The Latino gap is greatly influenced by *faminc*, *educ* and *age*.