

# Loan Originations and Defaults in the Mortgage Crisis: The Role of the Middle Class

## Internet Appendix

Manuel Adelino, Duke University

Antoinette Schoar, MIT and NBER

Felipe Severino, Dartmouth College

### Section IA.1 Sample definition

This paper focuses on 8,619 zip codes for which we have zip-code-level house prices from Zillow. Data for replicating Panels A and B of Table 2, as well as the IRS income quintiles as of 2002 are available at [https://dl.dropboxusercontent.com/u/23183326/A2S\\_2015\\_data.zip](https://dl.dropboxusercontent.com/u/23183326/A2S_2015_data.zip). These zip codes account for approximately 77% of the total purchase mortgage origination volume in HMDA.

We match census tracts from HMDA to zip codes using the Missouri Census Data Center bridge. This is a many-to-many match, and we rely on population weights to assign tracts to zip codes. The Missouri Census Data Center bridges of tracts to zip codes using population weights are obtained from <http://mcdc.missouri.edu/websas/geocorr90.shtml> for 1990 tract definitions (used in HMDA up to 2002) and <http://mcdc2.missouri.edu/websas/geocorr2k.html> for 2000 tract definitions (used in HMDA starting in 2003).

Below we show summary statistics for the full HMDA data set (Table IA.1). We have 27,385 zip codes in the HMDA data. In Table IA.2 we repeat the same specification as in Table 2 using all zip codes in the HMDA data.

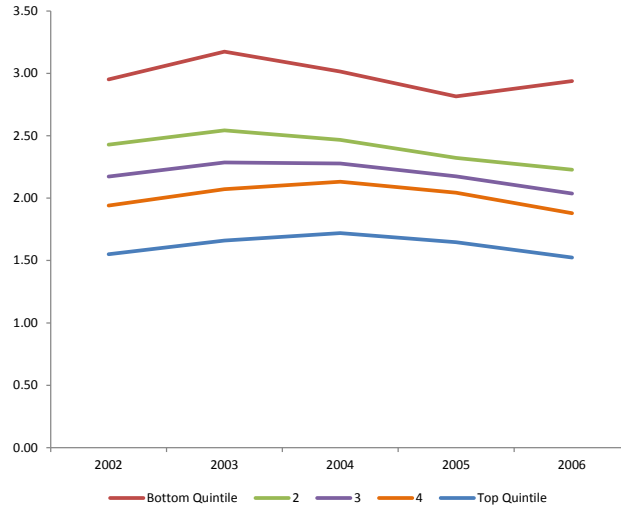
A comparison of Table IA.1 with Table 1 shows zip codes included in the Zillow sample are higher income and more densely populated than those in the full HMDA sample. The average mortgage size is \$155k in the Zillow sample, but only \$110k in the whole HMDA sample. Income growth is similar across the two samples, but purchase mortgage growth is generally higher when we include all zip codes in HMDA. This is mostly driven by higher growth rates for the number of mortgages by zip code.

Table IA.2 replicates the regressions in Table 2 with all zip codes in HMDA. The coefficient from the regression of the growth in mortgage origination on the growth in household income is positive and significant using the whole HMDA sample. The within estimator produces a positive coefficient, as does the between estimator. This suggests that, in the wider set of zip codes, the relation between growth in total mortgage credit and income is positive, independent of the use of county fixed effects.

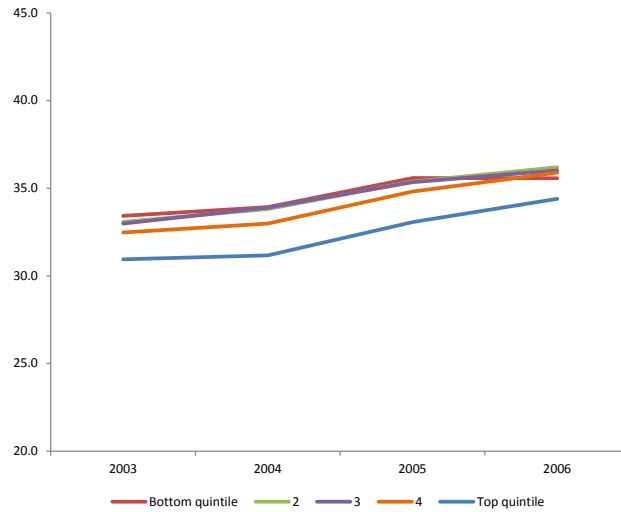
## Figure IA.1. Purchase mortgage DTI

Panel A shows average debt-to-income for individuals in each HMDA buyer income quintile. Debt-to-income is defined as the ratio of the mortgage balance at origination divided by the income of the applicant. Data are from HMDA, and the sample includes zip codes with nonmissing house price data from Zillow. Panel B shows debt-to-income for purchase mortgages according to a standard industry definition that calculates the ratio of recurring mortgage debt payments over monthly income. Income quintiles are based on IRS household income as of 2002. Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing house price data from Zillow.

### Panel A. DTI (individual buyer income, HMDA)



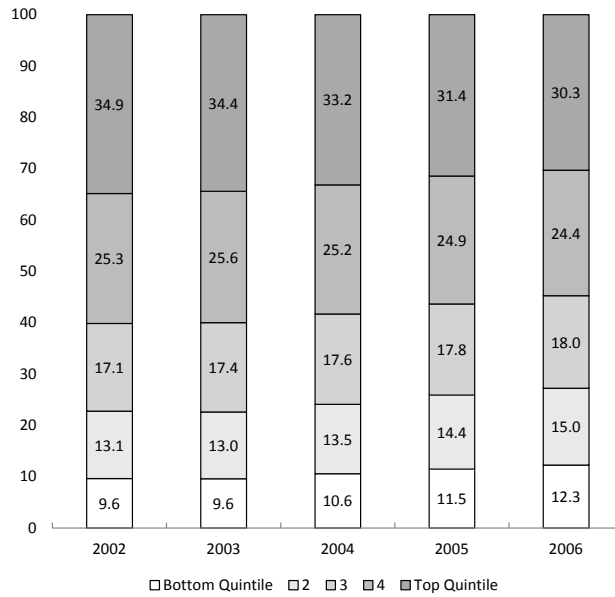
### Panel B. Mortgage payments as a percentage of income (quintiles based on IRS income as of 2002)



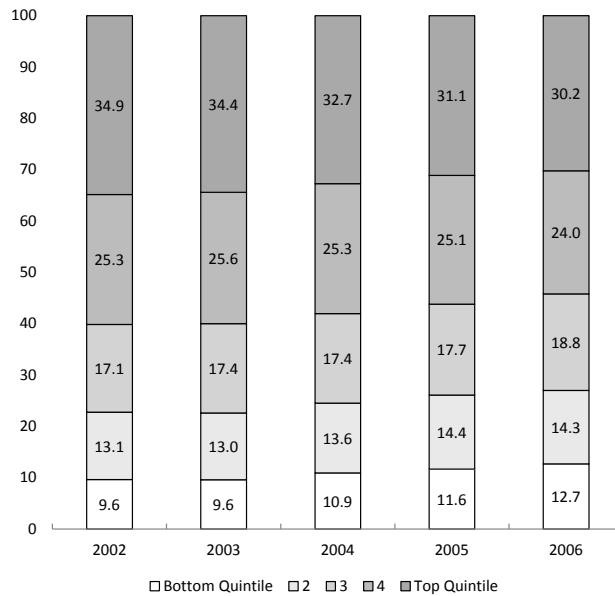
## Figure IA.2. Mortgage origination by income

This figure shows the fraction of total dollar volume of purchase mortgages in the HMDA data set originated by income quintile. In Panel A we form quintiles based on the household income from the IRS as of 2002 (i.e., the zip codes in each bin are fixed over time). In Panel B we resort zip codes based on household income as of each year (IRS data at the zip code level for 2003 are not available, so we use the ranking as of 2002 for both 2002 and 2003). The sample includes zip codes with nonmissing house price data from Zillow.

### Panel A. IRS Income quintiles as of 2002



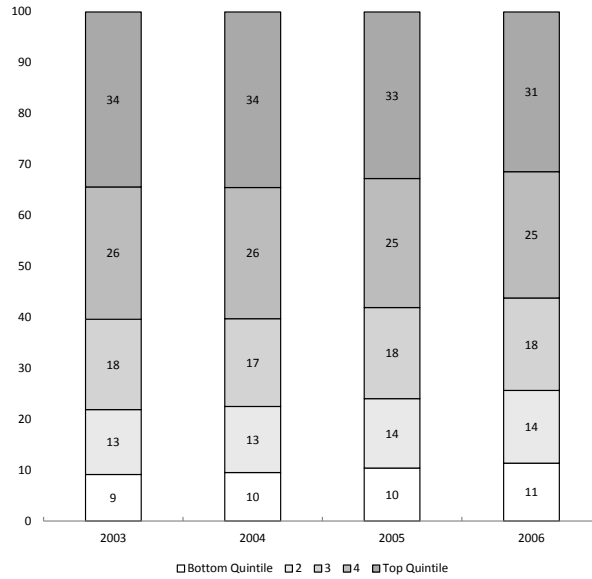
### Panel B. IRS Income quintiles as of each year



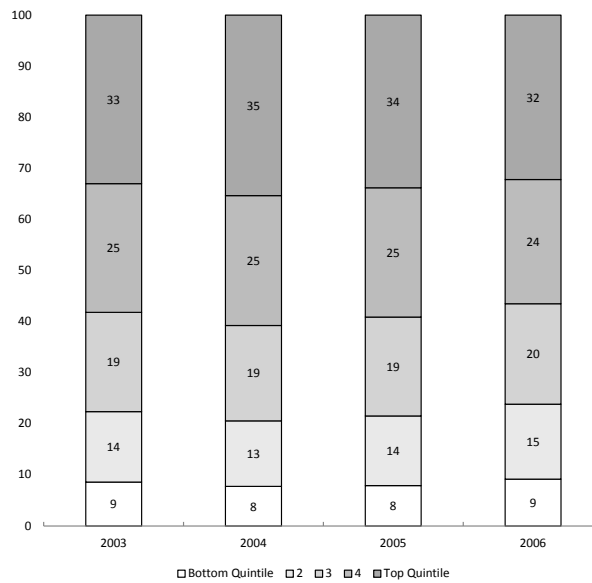
### Figure IA.3. Mortgage origination by income, LPS

This figure shows the fraction of total dollar volume of purchase mortgages originated in the 5% sample of the LPS data set. The sample includes zip codes with nonmissing Zillow house price data. In Panel A we form quintiles based on household income from the IRS as of 2002 (i.e., zip codes are fixed as of 2002, and cutoffs are the same as those given in Figure 1). In Panel B we form quintiles based on average buyer income from HMDA in the zip code as of 2002 (as of 2002 the zip code average buyer income cutoff for the bottom quintile is \$59k, the second quintile corresponds to \$69k, the third quintile corresponds to \$83k, and the fourth quintile corresponds to \$109k).

**Panel A. Origination (IRS household income quintiles as of 2002)**



**Panel C. Origination (HMDA buyer income quintiles as of 2002)**

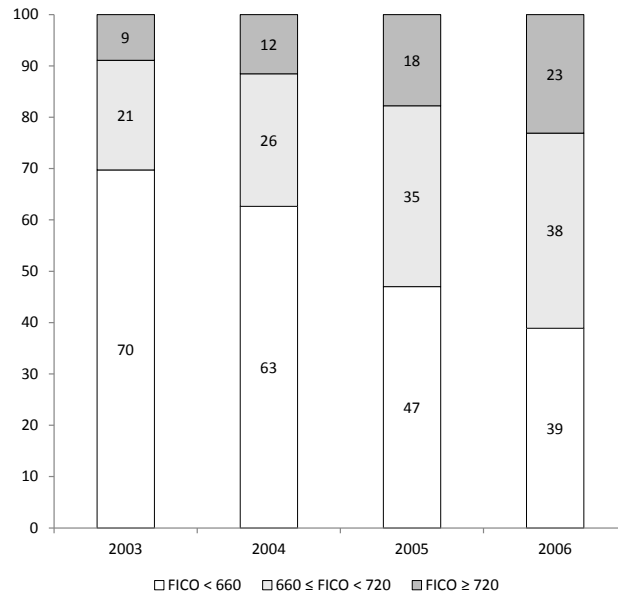


## Figure IA.4. Mortgage foreclosure by credit score

This figure shows the fraction of total dollar volume of purchase mortgages in foreclosure by cohort, split by income quintile. Data in Panel A are from the 5% sample of the LPS data set and in Panel B from Blackbox Logic, a data set of private-label securitized mortgages. The sample includes zip codes with nonmissing Zillow house price data. A mortgage is defined as being foreclosed on if it appears as in foreclosure or REO at any point during the three years after origination.

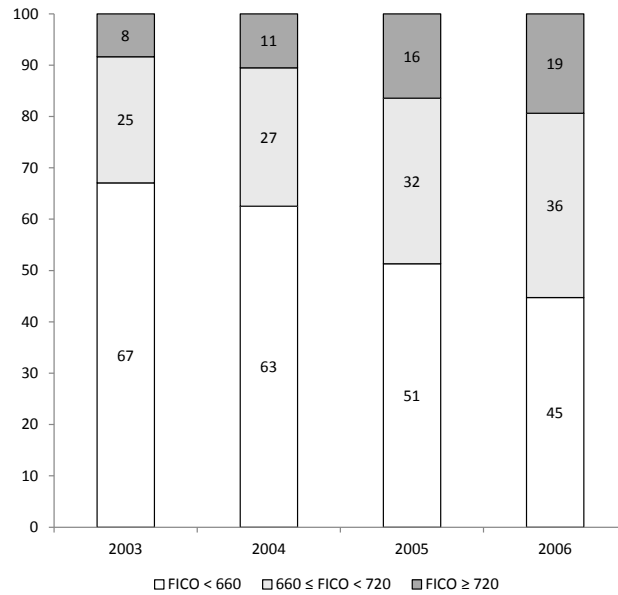
### Panel A. LPS data

---



### Panel B. Blackbox Logic data

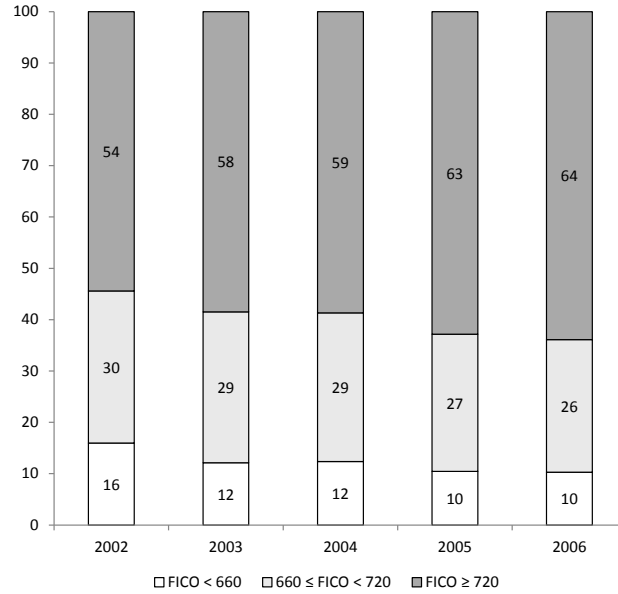
---



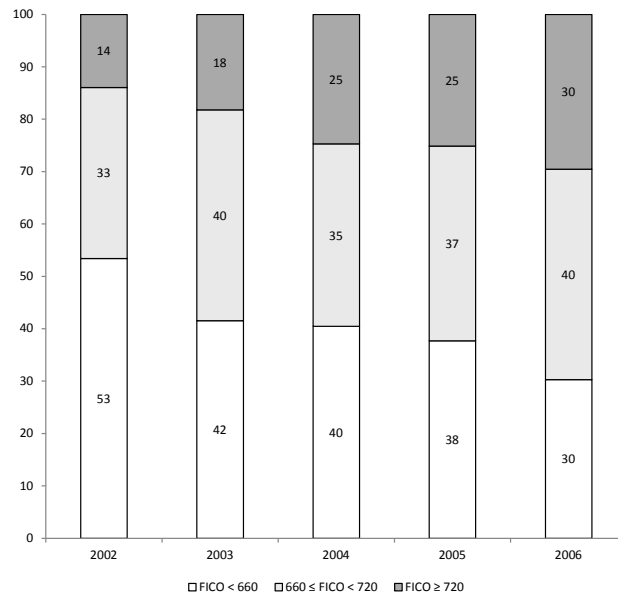
## Figure IA.5. Origination and delinquency by credit score, Freddie Mac data set

This figure shows the fraction of total dollar volume of purchase mortgages originated, as well as the total dollar volume of delinquent mortgages by cohort, split by credit score (FICO). A mortgage is defined as being delinquent if payments become more than 90 days past due (i.e., 90 days, 120 days or more, in foreclosure or REO) at any point during the three years after origination. Data come from the public Freddie Mac single-family home data set. The sample includes zip codes with nonmissing Zillow house price data.

### Panel A. Origination

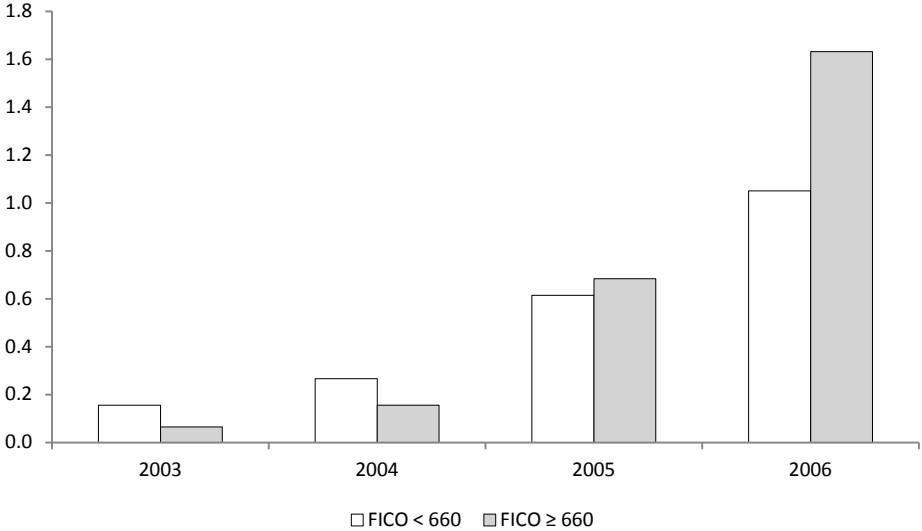


### Panel B. Delinquency



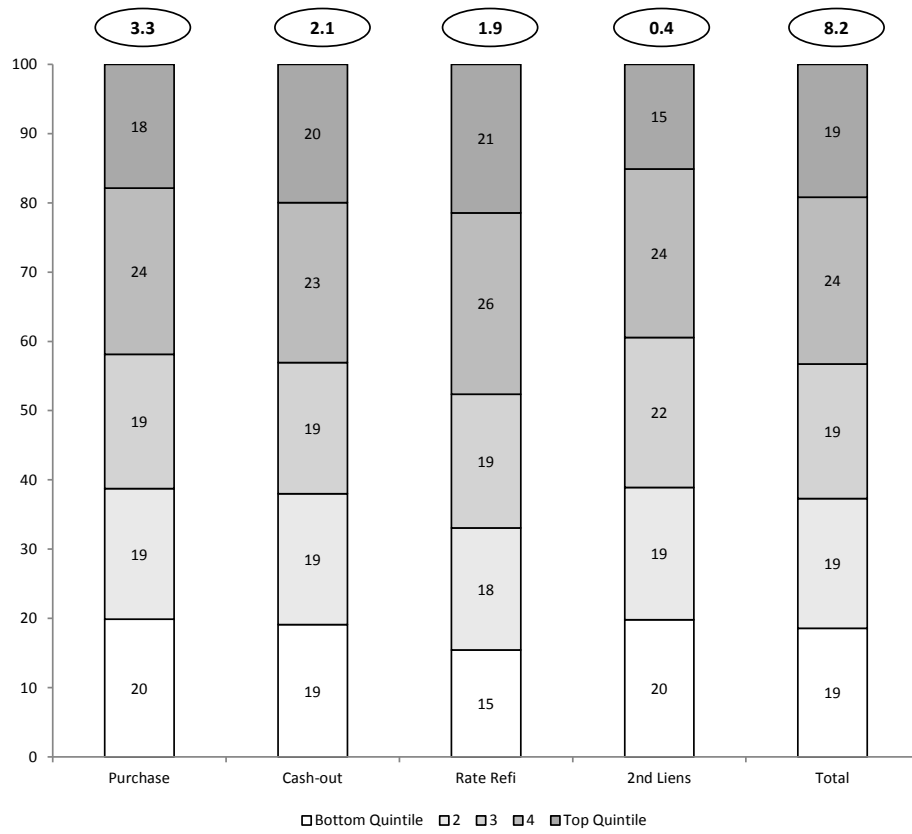
### Figure IA.6. Delinquency by credit score, dollar amounts

This figure shows the total dollar volume of delinquent purchase mortgages by cohort, split by credit score (FICO). A mortgage is defined as being delinquent if payments become more than 90 days past due (i.e., 90 days, 120 days or more, in foreclosure or REO) at any point during the three years after origination. Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing Zillow house price data. A FICO score of 660 corresponds to a widely used cutoff for subprime borrowers.



### Figure IA.7. Delinquency by product type and income, 2006

This figure shows the fraction of the total dollar value of mortgages that became 90 days delinquent or more at any time within three years after origination. The “Total” category includes mortgages that are unclassified in the data set. Total origination in billions of dollars in the LPS sample is shown above each bar. The sample includes zip codes with nonmissing Zillow house price data. Quintiles are based on household income from the IRS as of 2002, and cutoffs for each quintile are given in the notes to Figure 1.

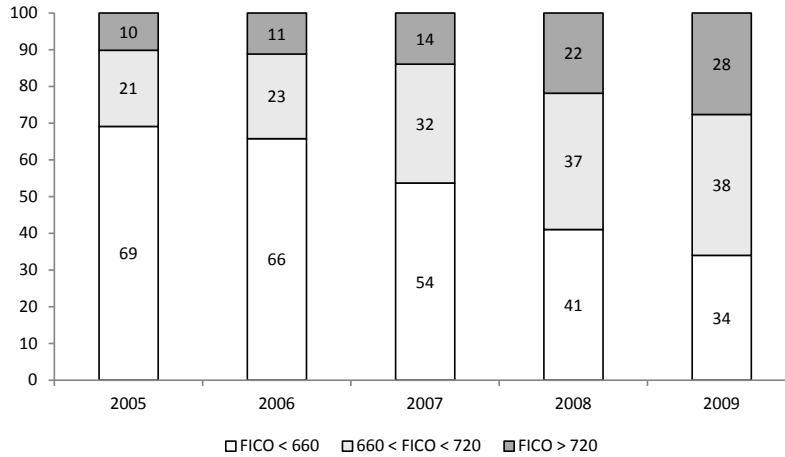




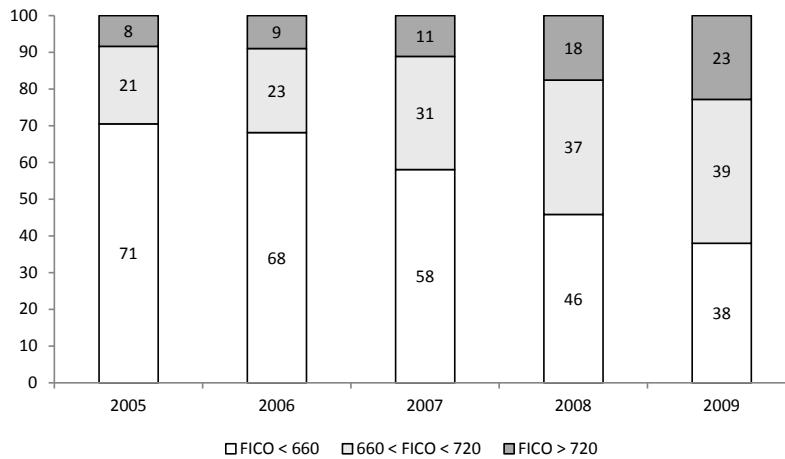
## Figure IA.8. Delinquency by calendar year and credit score

This figure shows the fraction of total dollar volume of delinquent purchase mortgages, as well as of all types of mortgages available in the LPS data set, split by credit score (FICO). Fractions are based on the total dollar amount of delinquent mortgages outstanding as of the last quarter of each year. A mortgage is defined as being delinquent if payments become more than 90 days past due (i.e., 90 days, 120 days or more, in foreclosure or REO) at any point during the three years after origination. Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing Zillow house price data. A FICO score of 660 corresponds to a widely used cutoff for subprime borrowers, and 720 is near the median FICO score of borrowers in the data (see Table 1 for additional summary statistics).

### Panel A. Purchase mortgages



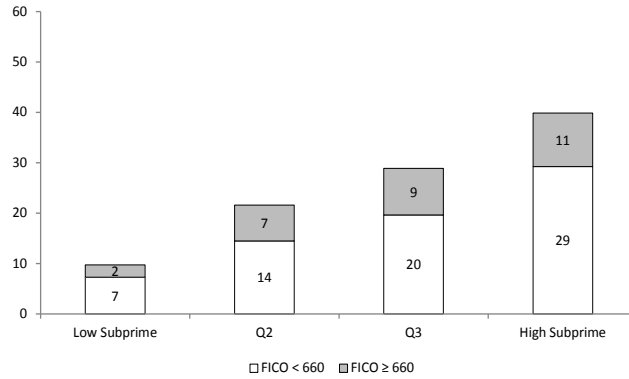
### Panel B. All mortgages



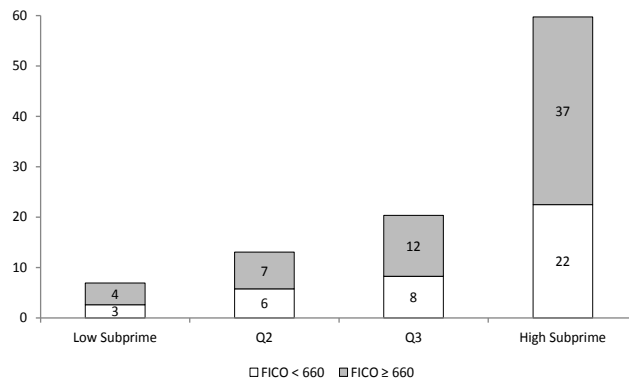
## Figure IA.9. Delinquency by proportion of subprime originations and credit score

This figure shows the fraction of the dollar volume of purchase mortgages more than 90 days delinquent at any point during the three years after origination for the 2003 and 2006 origination cohorts. Panels show splits by quartiles of subprime origination (defined as the percentage of loans made by lenders on the HUD subprime lender list), as well as by whether the borrower is above or below a credit score of 660 (a common FICO cutoff for subprime borrowers). In each panel fractions sum to 100 (the total amount of delinquent mortgages for each cohort), up to rounding error. Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing Zillow house price data.

### Panel A. 2003 Mortgage cohort



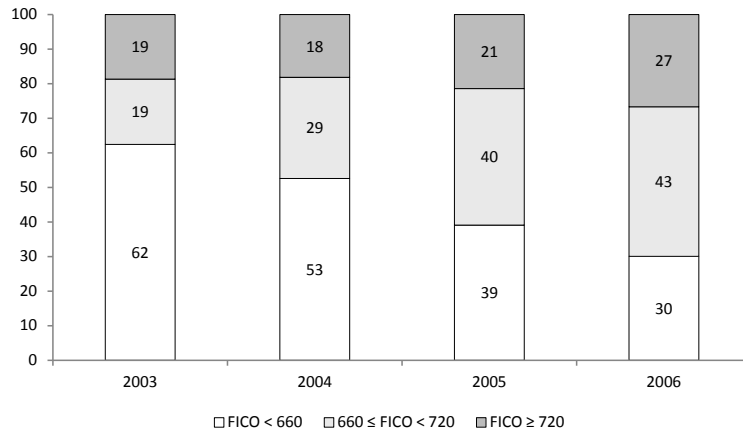
### Panel B. 2006 Mortgage cohort



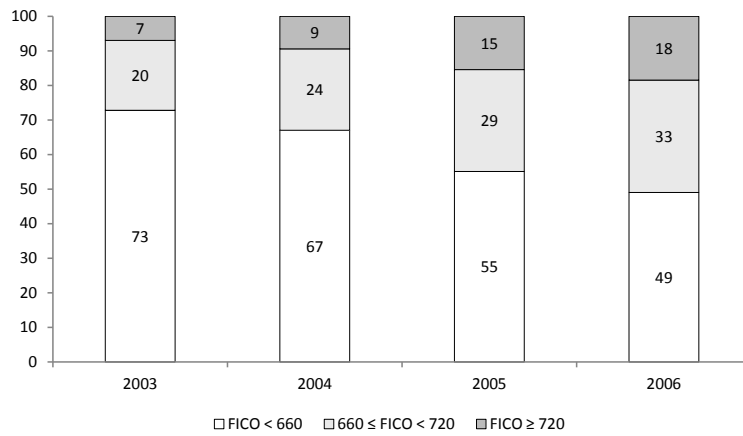
## Figure IA.10. Mortgage delinquency in recourse and non-recourse states

This figure shows the fraction of total dollar volume delinquent purchase mortgages by cohort, split by recourse and non-recourse states, as well as credit scores. Non-recourse states are AK, AZ, CA, HI, MN, MT, ND, OK, OR, and WA. A mortgage is defined as being delinquent if payments become more than 90 days past due (i.e., 90 days, 120 days or more, in foreclosure or REO) at any point during the three years after origination. A FICO score of 660 corresponds to a widely used cutoff for subprime borrowers, and 720 is near the median FICO score of borrowers in the data (see Table 1 for additional summary statistics). Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing Zillow house price data.

### Panel A. Non-recourse states



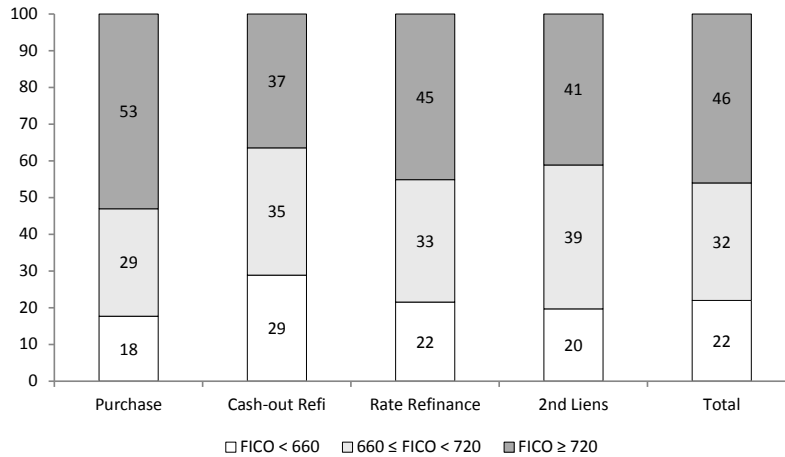
### Panel B. Recourse states



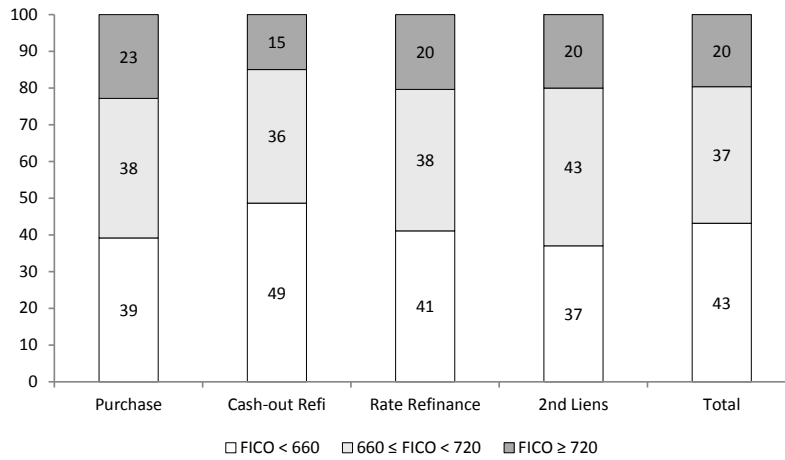
## Figure IA.11. Origination and delinquency by credit score, all mortgage types

This figure shows the fraction of total dollar volume of originations, as well as the total dollar volume of delinquent mortgages, of all types available in the LPS data set, split by credit score (FICO). A mortgage is defined as being delinquent if payments become more than 90 days past due (i.e., 90 days, 120 days or more, in foreclosure or REO) at any point during the three years after origination. A FICO score of 660 corresponds to a widely used cutoff for subprime borrowers and 720 is close to the median FICO score of borrowers in the data (see Table 1 for additional summary statistics). Data are from the 5% sample of the LPS data set, and the sample includes zip codes with nonmissing Zillow house price data.

### Panel A. Origination, 2006 cohort



### Panel B. Delinquency, 2006 cohort



**Table IA.1. Summary statistics for full HMDA sample**

This table reports summary statistics for all zip codes in the HMDA sample. Column 1 shows the pooled summary statistics. Columns 2–4 show the summary statistics by household income as of 2002 divided into the highest quartile (Column 2), the middle two quartiles (Column 3), and the lowest quartile (Column 4). For each variable we show the average and standard deviation (in parentheses). *IRS Household Income* is the average adjusted gross household income by zip code from the IRS. *HMDA Buyer Income* is the average applicant income by zip code from HMDA. *Average Purchase Mortgage Size* is the average size of purchase mortgages originated by zip code. *Number of mortgages originated per 100 residents* is the average number of purchase mortgages originated per 100 residents by zip code. *Debt to income* is the average ratio of the mortgage balance at the time of origination divided by the buyer income from HMDA.

	Whole	Zip code household income, 2002		
	sample	High	Middle	Low
	N =27385	N =6936	N =14126	N =6323
IRS Household Income, 2002, '000s	39.41 (24.68)	63.57 (39.10)	34.16 (3.65)	24.63 (3.41)
HMDA Buyer Income, 2002, '000s	74.21 (56.34)	111.06 (84.89)	64.11 (34.21)	56.33 (34.54)
Average Purchase Mortgage Size, 2002, '000s	110.47 (70.44)	182.26 (93.37)	93.10 (36.33)	70.53 (32.51)
Number of mortgages originated per 100 residents, 2002, purchase mortgages only	1.91 (6.08)	3.06 (8.49)	1.69 (5.73)	1.14 (2.39)
Debt to Income, 2002	1.83 (0.48)	2.16 (0.38)	1.79 (0.42)	1.55 (0.48)
Growth of IRS Household Income, 2002-2006, Annualized	0.043 (0.031)	0.051 (0.035)	0.039 (0.025)	0.045 (0.036)
Growth of HMDA Buyer Income, 2002-2006, Annualized	0.058 (0.072)	0.062 (0.064)	0.055 (0.065)	0.059 (0.090)
Growth in Total Purchase Mortgage Origination, 2002-2006, Annualized	0.164 (0.210)	0.110 (0.166)	0.172 (0.197)	0.204 (0.261)
Growth in Average Purchase Mortgage Size, 2002-2006, Annualized	0.066 (0.063)	0.065 (0.054)	0.063 (0.058)	0.076 (0.081)
Growth in Number of Purchase Mortgages, 2002-2006, Annualized	0.092 (0.180)	0.046 (0.150)	0.104 (0.169)	0.117 (0.220)

**Table IA.2. Mortgage origination and income, full HMDA sample**

This table shows regressions of growth in total purchase mortgage origination, the average purchase mortgage size, and the number of purchase mortgages originated at the zip code level on the growth rate of household income (from the IRS). For each measure we report the pooled OLS regression, a regression with county fixed effects, and the between county estimator. Growth rates are annualized and computed between 2002 and 2006. IRS data are available for 2002, 2004, 2005, and 2006. The sample includes all zip codes in the HMDA data set. Standard errors are clustered by county (shown in parentheses). \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

Estimator:	Total Purchase Mortgage			Average Mortgage Size			Number of Mortgages		
	OLS	Within	Between	OLS	Within	Between	OLS	Within	Between
Growth IRS Household Income	0.616*** (0.034)	0.554*** (0.032)	0.721*** (0.061)	0.446*** (0.014)	0.340*** (0.011)	0.472*** (0.027)	0.141*** (0.031)	0.191*** (0.027)	0.183*** (0.060)
County FE	-	Y	-	-	Y	-	-	Y	-
Number of observations	27,385	27,385	27,385	27,385	27,385	27,385	27,385	27,385	27,385
R2	0.04	0.33	0.08	0.25	0.61	0.24	0.00	0.29	0.01

**Table IA.3. Purchase and cash-out refinancing mortgages (LPS) and income**

This table shows regressions of the logarithm of mortgage size at the individual level on the logarithm of average household income in the zip code (inferred using zip code household income from the IRS). The unit of observation is an individual loan in the LPS data set originated between 2002 and 2006 in zip codes with nonmissing Zillow house price data. The first four columns include only purchase mortgages, and the last four include only cash-out financing mortgages. In Columns 2, 4, and 6 the income variable is interacted with dummy variables for each year in the sample. Standard errors are clustered by county (shown in parentheses). \*, \*\*, \*\*\* indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	Purchase				Cash-out Refinancing			
Ln(IRS Household Income)	0.623*** (0.014)	0.709*** (0.021)	0.454*** (0.047)	0.638*** (0.064)	0.550*** (0.016)	0.576*** (0.076)	0.284*** (0.060)	0.349*** (0.124)
Ln(IRS Household Income) x Year 2003		-0.082*** (0.013)		-0.058*** (0.013)	-0.046 (0.074)		-0.058 (0.088)	
Ln(IRS Household Income) x Year 2004		-0.078*** (0.016)		-0.049*** (0.017)	0.003 (0.080)		-0.009 (0.095)	
Ln(IRS Household Income) x Year 2005		-0.069*** (0.018)		-0.043** (0.020)	-0.003 (0.078)		-0.004 (0.094)	
Ln(IRS Household Income) x Year 2006		-0.124*** (0.019)		-0.108*** (0.021)	-0.054 (0.079)		-0.048 (0.096)	
Year FE and county FE	Y	Y	N	N	Y	Y	N	N
Year FE and zip code FE	N	N	Y	Y	N	N	Y	Y
Number of observations	272,077	272,077	272,077	272,077	108,097	108,097	108,097	108,097
R2	0.58	0.58	0.63	0.63	0.50	0.50	0.57	0.57