A Internet Appendix

A.1 Comparisons of Mortgage Debt, Income, and Debt Growth

Figure A.1 compares two aggregations of individual-level Equifax mortgage-debt balances from the New York Fed Consumer Credit Panel. The level of aggregation is either the state or the county. The horizontal axis of each panel measures our aggregation of debt for the given geographical unit, while the vertical axis measure aggregates that are calculated and published by the New York Fed itself. In all panels, the dots lie along 45-degree lines, giving us confidence that we are aggregating up to the zip-code level correctly when we construct our main cross-sectional dataset.

In the text, Figure 3 compared the aggregate number of returns from the zip code-level data with the aggregate number of returns published by IRS; the latter series omits any return filed for the sole purpose of receiving an economic-stimulus payment. In most years, the total number of returns in the zip code-level data is smaller than the IRS's published total, in part because of the suppression rules. But as Figure 3 shows, in 2007 the zip code-level data imply many more returns, because these data include returns filed for the sole purpose of receiving stimulus checks. Figure A.2 compares aggregates of IRS income data rather than numbers of returns. The blue lines in this figure are national aggregates of either salary and wage income (top panel) or AGI (bottom panel) published by IRS. The red dots are aggregated data from the zip code-level IRS dataset that we use in the paper. In both panels, the published aggregate is larger than the zip code-level aggregate, probably because of the suppression rules that IRS applies to the zip code-level dataset before they release it. Even so, the two income aggregates follow similar time-series patterns—even in 2007, when the number of returns filed spikes up.

Figure A.3 compares estimates of the aggregate stock of mortgage debt from the Flow of Funds, the Equifax dataset, and the SCF in those years for which the SCF data are available. The Equifax totals are close to, but somewhat smaller than, the SCF and Flow-of-Funds totals. Yet our Equifax debt totals are essentially identical to some unreported Equifax totals calculated by Brown et al. (2015), who compare Equifax data with the SCF along a number of dimensions.[59] Two SCF aggregates are presented. The first comes from Henriques and Hsu (2014), who compare various SCF aggregates to their Flow of Funds counterparts. Even though Flow of Funds data are typically constructed from administrative records supplied by financial institutions and government agencies, rather than from surveys,

[59]Specifically, in billions of 2010 dollars, Brown et al. (2015) estimate total mortgage debt in Equifax in 2004, 2007, and 2010 to be \$7,631, \$10,034, and \$9,282, respectively. Our Equifax totals expressed in the same units and years are \$7,741, \$9,728, and \$9,074. In addition, unreported work shows that our totals are close to those reported Bhutta (2015), which also analyzes mortgage debt in the New York Fed Consumer Credit Panel.

Henriques and Hsu (2014) show that most balance-sheet measures in the SCF are close to the corresponding Flow of Funds estimates. This comparability is particularly true for mortgage debt, a pattern the authors attribute to the clarity of the mortgage debt concept and the stability of mortgage data collection procedures in both the SCF and the Flow of Funds over time. Figure A.3 replicates the close correspondence between mortgage debt in the Flow of Funds and Henriques and Hsu's SCF measure. Gratifyingly, the figure also shows that our SCF aggregates, based on the public-use summary SCF datasets, are essentially identical to Henriques and Hsu's, with the small differences between them probably resulting from the fact that we use the public-use version of the data. Note that comparability of the SCF data to the mortgage measure in the Flow of Funds requires the use of all mortgage data available, including HELOCs. This is why we include HELOCs and other types of secondary mortgages when using either the Equifax dataset or the SCF.[60]

A.2 AGI vs. Wages and Salaries and 2006 vs. 2007

Throughout the paper we use salary and wages to measure income, because this type of income is most likely to be the focus of lenders when they evaluate mortgage applications. An alternative choice would be to use AGI (for zip code-level analysis) and total income (for SCF analysis). Figure A.4 shows that our main results hold even when this alternative choice is made. The two left panels use zip code-level data from Equifax, with quintiles calculated using AGI per tax return. The two right panels use data from the SCF, with the quintiles calculated using the SCF's measure of total income. The top panels show the same similarity in debt evident in Figure 2 in the main text in both the zip code-level data and the SCF. The lower panels replicate the finding that, because mortgage debt rises with income in the cross section, equal debt-growth rates imply very large dollar amounts of new debt for the richest borrowers.

The main text also uses 2006 rather than 2007 as the last year of the mortgage boom when performing zip code-level analysis. (The ending-year issue is not relevant for the SCF.) This choice is necessitated by the spike in tax filing in 2007 illustrated in the text by Figure 3 and discussed earlier in this appendix. Recall that Figure A.2 implied that the additional filers had very low incomes, because their tax returns had little effect on 2007 levels of total AGI or of wages and salaries. Further evidence that the extra filers had low incomes appears in Figure A.5. This figure shows that zip code-level growth in the number of IRS returns filed in 2007 is not only much greater than in other years, but that 2007 growth covaries

[60] Recent work by Amromin and McGranahan (2015) and Amromin, McGranahan, and Schanzenbach (2015) also uses the Equifax dataset but splits mortgage debt into non–HELOC mortgage debt and HELOCs. Although these papers do not emphasize the point, they also find broadly similar growth rates of mortgage balances across the income distribution, even when HELOC balances are excluded.

negatively and monotonically with zip code-level income. As with the choice of income definition, however, the choice of ending year has little effect on the main results. Figure A.6 shows that using 2007 as the end of the boom for the zip code-level distributions generates the same patterns seen in earlier figures.

A.3 Individual-Level IRS Data from the Tax Model Files

The analysis in the main text uses publicly available income data from the IRS that has been aggregated to the zip-code level. However, as noted in the text, the IRS also produces a public-use sample of anonymized individual-level tax returns, often referred to as the Tax Model Files. These files are well suited for studying the U.S. income distribution due to their large sample sizes (samples for recent years include around 150,000 individual tax records) as well as their excellent coverage of the highest-income tax filers. Recently, Saez and Zucman (2016) used the Tax Model Files to study changes in the U.S. wealth distribution over time, by capitalizing flows of reported income and deductions into stocks of wealth and debt. As part of that study, the authors constructed individual-level estimates of housing assets by capitalizing property tax payments in a way that forced implied totals to be consistent with national aggregates. Similarly, mortgage debt was netted out of housing wealth by capitalizing mortgage-interest deductions.

The Tax Model Files are much larger than the SCF and are available annually rather than every three years. Even so, the SCF provides better coverage of housing wealth and debt throughout the whole of the income distribution, in large part because lower-income filers are less likely to itemize their property-tax and mortgage-interest deductions. For homeowners who take the standard deduction rather than itemize, there is no way to capitalize their property-tax and mortgage-interest payments into stocks of housing wealth or mortgage debt. When the focus is on the stock of wealth at the top of the income distribution, as in Saez and Zucman (2016), low itemization rates in the middle and bottom of the income distribution are not a significant concern. But when the focus is on mortgage debt throughout the income distribution, as in this paper, the lack of housing-related information for the mass of tax filers is a serious drawback.

For an estimate of how serious this drawback is likely to be, we turn first to Poterba and Sinai (2008), a paper that assessed the likely incidence of various housing-related tax provisions for homeowners of varying ages and incomes. As part of that study, the authors used the NBER TAXSIM federal and state income-tax calculators to estimate how many homeowners in the 2004 SCF would be better off taking the standard deduction as opposed to itemizing their deductions, and thus becoming eligible for tax benefits such as the mortgage-interest deduction. The Poterba-Sinai results are displayed in the first two columns of Table A.1. The authors found that only 23.4 percent of homeowners with under $40,000 in annual income should itemize their deductions. The predicted itemization rate for homeowners with \$125,000–\$250,000 of income was 98.4 percent, while the comparable figure for homeowners with incomes above \$250,000 was 99.9 percent.[61] These results imply that Saez and Zucman (2016) have a near-complete sample of property-tax and mortgage-interest payments at the top of the income distribution. But applying the capitalization method to the lower parts of the income distribution would miss many homeowners.

To further illustrate the inverse relationship between income and itemization rates among homeowners, we make a similar calculation using the IRS zip-code level data. The 2007 data include information on the number of tax returns in each zip code that reported mortgage interest payments. Using the Equifax data to estimate the number of households with a mortgage in each zip code, we can then calculate the fraction of households with a mortgage who deducted their mortgage interest paid. We do this for five income quintiles of zip codes, using either average AGI or average salary and wages to construct quintiles. The results are reported in the remaining columns of Table A.1. Consistent with the predictions of Poterba and Sinai (2008), we find that a higher fraction of mortgaged households in high-income zip codes deduct mortgage interest payments. Because our calculation is done at the zip-code and not the household level, our results are not as extreme as those of Poterba and Sinai (2008). We find that around 45 percent of mortgaged households in the lowest-income quintile itemize their mortgages, compared to nearly 75 percent in the highest-income quintile.

The Poterba-Sinai calculation implies that 63.1 percent of all homeowners in the 2004 SCF should have itemized their deductions. That figure is quite close to the overall itemization rate of 63.8 percent that we find using the 2007 zip-code level data. Interestingly, Poterba and Sinai also report that their prediction is close to the fraction of 2004 SCF homeowners who reported they did in fact itemize on their most-recent return (63.3 percent). However, at a more disaggregated level, Poterba and Sinai find large differences in predicted vs. self-reported itemization rates among low- and middle-income homeowners. Among homeowners making less than \$125,000 per year, the self-reported itemization rate for those headed by individuals younger than 35 was 20 percentage points lower than predicted. For homeowners over 65, the self-reported itemization rate was about 20 percentage points higher than predicted. For our purposes, the implication of this finding is that it would be difficult to impute itemization rates among low- and middle-income filers based on income levels.

Finally, we note that an additional problem with using the capitalization method to

[61]The authors define household income as “adjusted gross income plus the following items: income from nontaxable investments, an estimate of employer contributions for FICA, payments from unemployment insurance and workers compensation, gross social security income, and any alternative minimum tax (AMT) preference items that can be estimated from the SCF” (p. 84).

estimate individual-level debt levels for low- and middle-income taxpayers is that this method assumes identical capitalization factors across income groups. When measuring stocks of wealth, as in Saez and Zucman (2016), the assumption of equal capitalization factors makes sense. Two individuals holding an identical bond will receive identical interest payments, even though one individual may be richer than the other. Because the interest rate on the bond is essentially the inverse of the capitalization factor, the capitalization factor used to infer stocks of bond holdings from flows of bond interest should therefore be identical across income groups. However, mortgage interest rates *paid* by two individuals are likely to differ with income.[62] As a result, assuming an identical capitalization factor to back out stocks of mortgage debt from flows of mortgage interest is less tenable when households have very different incomes.

A.4 Relationship to Kumhof, Rancière, and Winant (2015)

In a stimulating paper on the potential relationship between income inequality and financial crises, Kumhof, Rancière, and Winant (2015, henceforth KRW) use the SCF to motivate a general equilibrium model of lending across income groups. The model is designed to show how rising income inequality in the 1920s and the latter part of the 20th century could have contributed to the Great Depression and the Great Recession, respectively. In the model, households at the top of the income distribution lend to the rest of the population, so that these borrowers can maintain their consumption levels amid declining relative incomes. At some point, the lower-income households make a rational decision to default on their consumption loans, a choice that gives them financial relief but that also triggers a financial crisis. To motivate the model, the first part of the paper uses SCF data to plot the total debt-to-income ratio (DTI) of the top 5 percent of households along with that of the bottom 95 percent.[63] The DTI of the bottom 95 percent trends up from 1983 through 2007. The DTI of the richest 5 percent has no discernible trend from 1983 onward, although it does increase somewhat from 2001 to 2007. The data are thus consistent with a broad prediction of the KRW model: before the Great Recession, the debt burden of the bottom 95 percent of households grew relative to the debt burden of the households at the top.

The DTI plot motivates a model of *total* debt, not *mortgage* debt, but KRW's analysis of debt for different income categories has some similarities to ours, so comparison of the two approaches is informative. As a first step in this comparison, the two left-hand panels of Figure A.7 plot mortgage debt levels and shares across 20 bins of the income distribution, rather than the five bins used in Figure 2. The use of 20 rather than five categories allows an examination of mortgage debt of the richest 5 percent of households, one of the two income

[62] Mortgage interest rates are also likely to differ with the age of mortgages and other factors.
[63] See Panel A of Figure 2, p. 1222.

groups in the KRW graph. The additional bins do not change the main message of Figure 2, as debt shares remain relatively stable and the richest categories take out the most debt in dollar terms. The share of mortgage debt of the very top bin drops somewhat, but the top 5 percent of households takes out \$618 billion in new mortgage debt during the boom, more than the \$554 billion of the bottom 40 percent combined.

Like the income measure featured in the main results of this paper, income in Figure A.7 is defined as salary and wage income, whereas KRW use total income. Table A.2 therefore presents a series of mortgage-debt and income statistics for the bottom 95 percent and top 5 percent of the total-income distribution. The first two rows of the table emphasize the disproportionate representation of high-income households in outstanding mortgage debt. By definition, the top 5 percent of households always account for one-twentieth of all households, but their share of aggregate mortgage debt was slightly higher than one-fifth when the mortgage boom began in 2001. The next row shows that in dollar terms, the average household in the top 5 percent started the boom with an outstanding mortgage-debt balance of just over \$200,000, which rose to just under \$350,000 by 2007. In dollar terms, the richest group accounts for 20 percent of the aggregate dollar-value increase of additional mortgage debt from 2001 to 2007 (row 4). Because this percentage is slightly below the top group's 23 percent share of mortgage debt at the start of the boom (row 2), the top group's share of aggregate mortgage debt falls—to 21 percent—in 2007. The use of total income rather than salary and wage income therefore results in a small drop in the share of mortgage debt held by the top 5 percent, similar to the decline seen using salary and wage income in Figure A.7.

The KRW model studies how the debt burden of the bottom 95 percent affects its incentives to default and thus the health of the financial system, and it measures leverage using debt-to-income ratios (DTIs). Rows 5 through 8 of Table A.2 characterize the evolution of total household income for the two income categories. The first of these rows shows that average income of the high-income group is close to \$500,000 in 2001, rising to around \$630,000 six years later. The implied rate of income growth is larger than that experienced by the lower-income group, so the share of income earned by the top 5 percent rises, from 35 percent in 2001 to 37 percent in 2007. Row 7 shows that the mortgage DTI of the bottom 95 percent rises from 0.77 to 1.22, while that of the higher-income group increases from 0.41 to 0.55. Thus, the mortgage data are consistent with a visual implication of KRW's graph, as the absolute change in the DTI of the bottom 95 percent is larger than the absolute change of the top 5 percent from 2001 to 2007. When debt is defined as mortgage debt alone, as in Table A.2, the absolute increases in DTIs are 0.45 and 0.14, respectively (row 8). Differences in absolute DTI in the KRW figure appear even larger than this. This is almost certainly because the KRW paper plots total DTIs rather than mortgage DTIs, and KRW find that non-mortgage debt rises more quickly for the lower-income group.[64]

Absolute differences in DTIs could well be useful for motivating a general equilibrium model of indebtedness, but they are less informative about changes in the distribution of mortgage debt with respect to income.[65] If income levels of the two groups had not changed from 2001 to 2007, equal percentage increases in mortgage debt for the two groups would have generated equal percentage changes in DTIs. However, because the bottom 95 percent began the boom with a higher DTI than the top 5 percent (0.77 vs. 0.41), equal percentage changes in debt would have generated a larger absolute change in the DTI of the lower-income group. Additionally, the faster income growth experienced by the richest group from 2001 to 2007 would have reduced its relative DTI growth in both absolute and relative terms. When we consider *log changes* rather than *percentage changes*, the effects of changes in either debt or income on DTIs are exact (notwithstanding rounding). Row 9 of Table A.2 shows that the log changes in the two DTIs are 0.46 and 0.30, which are more similar than the absolute changes in the previous row. The difference of 0.16 in the log change is accounted for nearly equally by somewhat faster debt growth among the lower-income group (row 10) combined with somewhat slower income growth (row 11).

While our data are consistent with KRW's plot, we interpret the main reason behind the overall increase in debt differently than they do. Flow of Funds data indicate that mortgage debt grew much more quickly than non-mortgage debt during the early 2000s, so that the explanation of the overall debt boom during this period must involve mortgages.[66] Some mortgage debt undoubtedly supported consumption, in the spirit of KRW's model. But one-fifth of the aggregate mortgage-debt increase was accounted for by the richest 5 percent of households, a group that averaged nearly a half-million dollars in annual income at the start of the boom. This fact suggests that borrowing for investment, not consumption, was the true driver of debt accumulation. Along these lines, the right-most panels of Figure A.7 show that the disproportionate share of real estate *assets* held by the top 5 percent became even larger during the boom. Although these cross-sectional patterns suggest that investment motives were paramount during the boom, we recognize that a comprehensive comparison of consumption and investment motives must await future research.

[64] See Panel B of KRW's Figure 3 (p. 1222).

[65] While indebtedness is often measured with DTIs, using DTIs to study the evolution of debt in the cross-section is similar to using a less-flexible version of the benchmark log-log regression employed in the main text. The benchmark model regresses the natural log of mortgage debt for a cross-sectional unit on the log of its income and either a constant or city fixed effects. Working with DTIs constrains the coefficient that multiplies log income in this regression to equal one.

[66] Figure 1 shows that mortgage debt grew much faster than personal disposable income during the early 2000s, but we found in unreported work that non-mortgage debt grew about as fast as income during the boom.

A.5 Debt Distributions Disaggregated by Lien Types

In the text, we investigate debt patterns using all types of mortgage debt: first mortgages, second mortgages, and HELOCs. Figure A.8 disaggregates the analysis by lien type. For reference, we include as the upper left panel of this figure the overall debt distribution with respect to income that appeared as part of Figure 2 in the text. The top right panel of Figure A.8 shows the distribution of first-mortgage debt. Because the large majority of outstanding debt consists of first liens, it is not surprising that the first-lien distribution remains stable over time. The lower left panel presents distributions of closed-end second mortgages. Here, there is a pronounced change in the distribution, with high-income ZIP codes receiving much higher shares of second-mortgage debt in 2006 relative to 2001. The last panel shows distributions of HELOC debt. There is a slight tilt toward higher debt shares among richer quintiles, but this tilt is not as severe as in the previous panel. In any case, none of the panels in Figure A.8 indicates a significant increase in the share of debt held by low-income quintiles. Figure A.9 performs the same analysis using AGI rather than salaries and wages, with similar results.

A.6 Distributional Statistics for the SCF

The panels in Figure A.10 provide some formal statistics for the SCF distributions in the top left panel of Figure 4. The top left panel graphs the mean and median debt levels for each year of the SCF, the top right panel depicts the standard deviation, and the two bottom panels plot the inter-quartile range and the 90th–10th percentile differences, respectively.

A.7 Homeownership in the SCF

Figure A.11 depicts income coefficients from a regression of homeownership on income. This analysis is structured analogously to the mortgageship logits in Figure 7 in the text, and shows essentially the same patterns.

A.8 Identifying First-Time Mortgageship in Equifax

As mentioned in the main text, Equifax includes a variable that gives the age of the oldest mortgage on record for each individual. Using this variable, we identify individuals taking out their first mortgage. However, the information in this variable needs to be cross-verified with other variables in Equifax before it is usable. We code someone as taking out their first mortgage if the individual is born in 1950 or later,[67] and if the age of her oldest mortgage

[67] We limit our analysis to borrowers born in 1950 or later because for borrowers born earlier, it is much more likely that the first mortgage recorded in Equifax is not actually their first mortgage, because Equifax

is zero within one quarter on either end of originating a first-lien mortgage. In other words, an individual could have a first-mortgage originated in the quarter just prior or just after Equifax indicates that the age of a first mortgage has gone from “no account on file” to zero. For the hazard ratios in the main text, we are interested in the probability of an individual taking out their first mortgage, so we do not correct for joint mortgages. However, in the data check described in the next paragraph we divide the number of joint mortgages by two.

Figure A.12 plots our estimate of the share of all purchase mortgages going to people taking out their first mortgage, along with the estimate of the share of all purchase mortgages going to first-time homeowners from the National Association of Realtors (NAR) Annual Survey. It should be noted that first-time mortgageship and first-time homeownership are not necessarily the same: it is possible that someone inherited a home or bought a home with cash prior to taking out a first-ever mortgage. However, the two should be highly correlated. Our estimate is calculated using a 10 percent sample from Equifax and HMDA. The numerator is the number of first-time-ever mortgages originated in Equifax and the denominator is the number of owner-occupied purchase mortgages in HMDA. In comparison, NAR is a survey of over 100,000 home buyers with a less than 10 percent response rate, so their final survey results have under 10,000 home buyers and suffer from selection bias. Given that each of these two measures is imperfect, their similarity is remarkable.

A.9 Gross Flows Analysis without Area-Level Fixed Effects

The two panels of Figure A.13 present the estimated income effects from regressions that have either total originations (top panel) or total terminations (bottom panel) on the left-hand side. They are analogous to the right-hand-side panels of Figure 10, which are generated from regressions that also include CBSA $\times$ year fixed effects.

A.10 Effect of Subprime and Alt-A on Debt Growth: 2001–2007

Figure A.14 replicates the main lessons of Figure 15 for debt growth, using 2007 as the ending year of the mortgage boom. Like Figure 15, which ends the boom in 2006, the appendix figure shows that the use of subprime mortgages grew more in low-income areas. Growth in Alt-A and prime mortgage debt tended to be higher in richer ZIP codes. The end result is that mortgage debt grew at broadly similar rates throughout the income distribution, both within and across CBSAs.

did not computerize its records until after 1970.

A.11 Alternative Measures of Homeownership

Figure A.15 provides two estimates of the homeownership rate. The red line shows the standard homeownership rate, which is calculated by the Census Bureau as the share of occupied housing units that are occupied by owners. This rate is calculated using owner-occupied rates in the Current Population Survey/Housing Vacancy Survey. The blue line shows an alternative measure of homeownership suggested by Mian and Sufi (2016b), which is the total number of owner-occupied units divided by the population of adults aged 15 or older. The latter measure requires an estimate of the total number of owner-occupied units, which is periodically updated by Census based on changes in methodology or sample frames. Our estimate adjusts for the break in the estimate of owner-occupied units that occurs in 2000. This adjustment uses information on the total housing inventory in Table 953 of the 2004 Statistical Abstract of the United States.

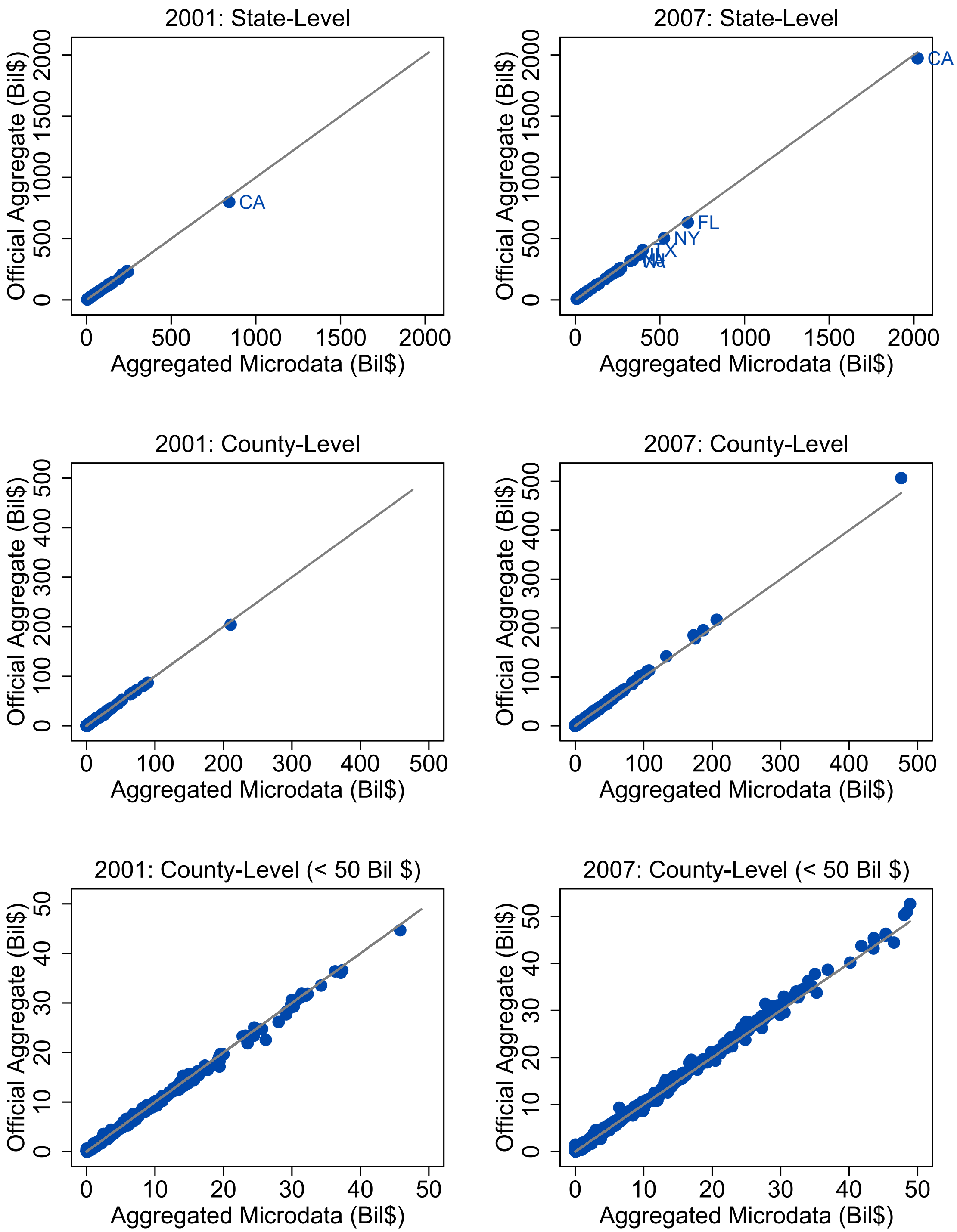


Figure A.1. Comparison of Aggregated Mortgage Debt Balances in the New York Fed Consumer Credit Panel. *Note:* Each of the panels above is a comparison of aggregated data from the microlevel records of the New York Fed Consumer Credit Panel. Aggregation along the horizontal axes was performed by the authors, while the vertical axes measure aggregates generated from the same dataset by the Federal Reserve Bank of New York. For the county-level data in the lower two rows, only counties with at least 10,000 consumers possessing credit records in 2010:Q4 are included. *Source:* New York Fed Consumer Credit Panel/Equifax.

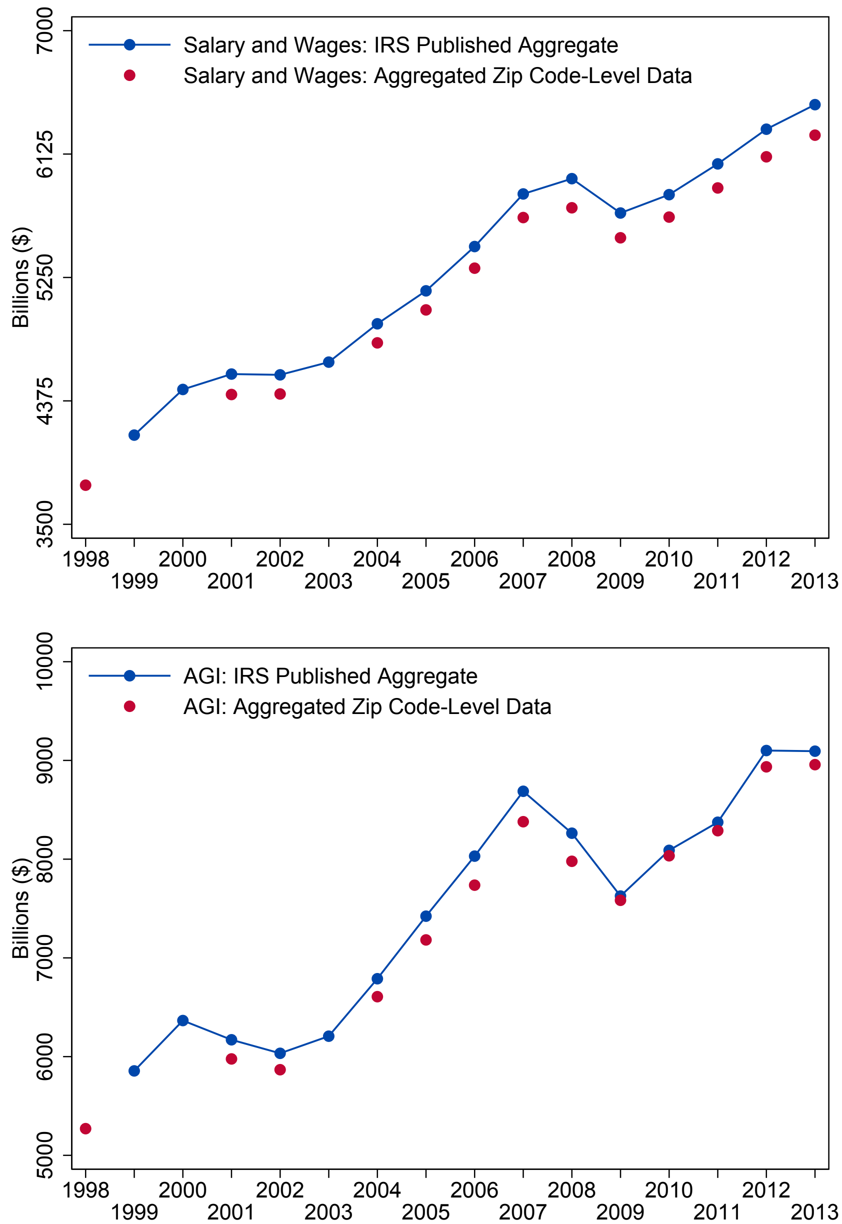


Figure A.2. Measures of Aggregate Salary and Wage Income and Adjusted Gross Income. *Note:* In each panel, the blue line depicts the given income aggregate as published by the IRS, and the red dots depict annual aggregates generated from the zip code-level IRS data. *Source:* Internal Revenue Service, Statistics of Income Historical Table 1 (available at `https://www.irs.gov/uac/SOI-Tax-Stats-Historical-Table-1`).

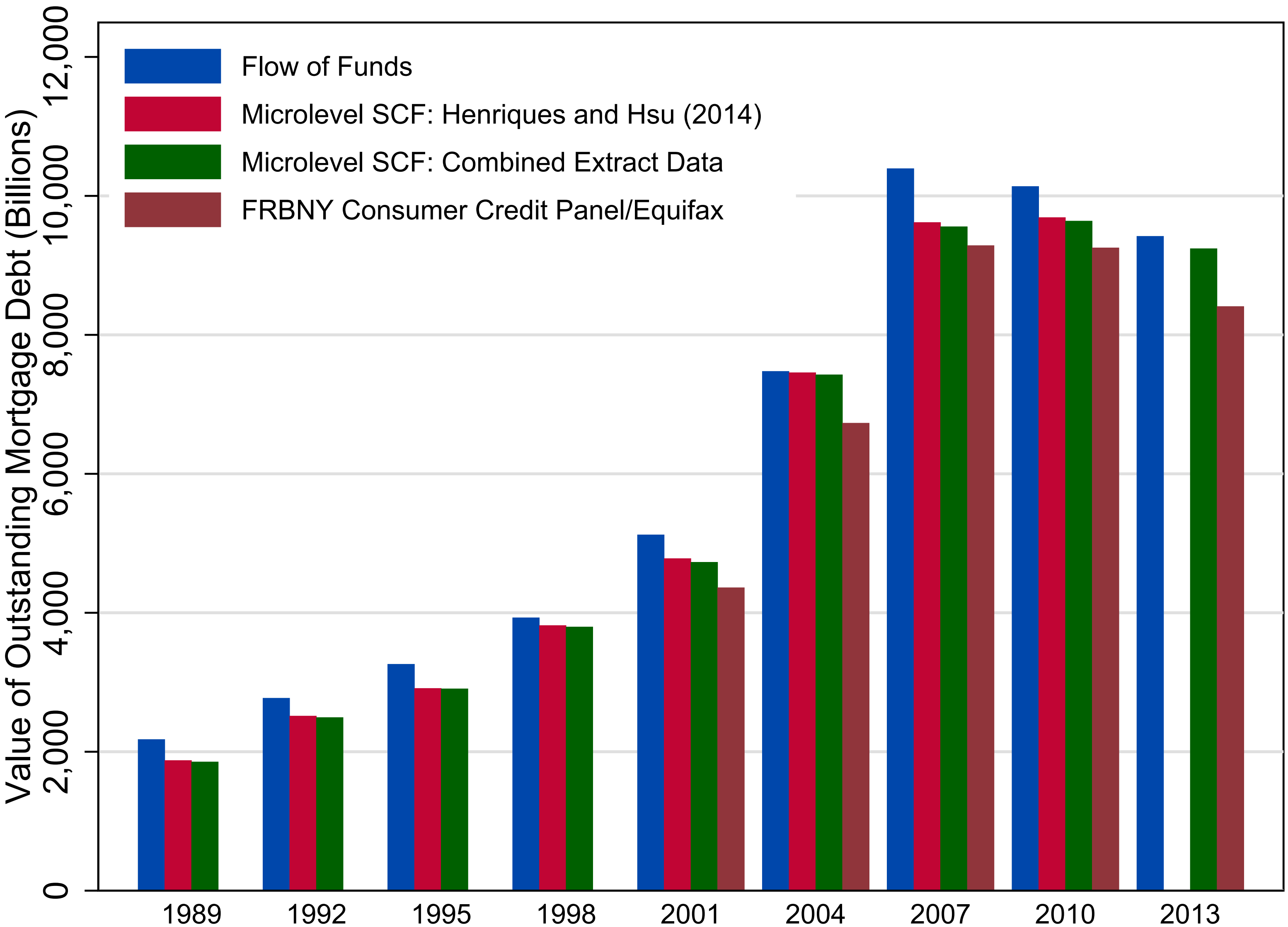


Figure A.3. Alternative Measures of Aggregate U.S. Mortgage Debt. *Source:* Board of Governors of the Federal Reserve System (for Flow of Funds); Table 9.1 (p. 250) of Henriques and Hsu (2014); authors' calculations using the Combined Extract Data of the Survey of Consumer Finances; and authors' calculations using the NY Fed Consumer Credit Panel/Equifax.

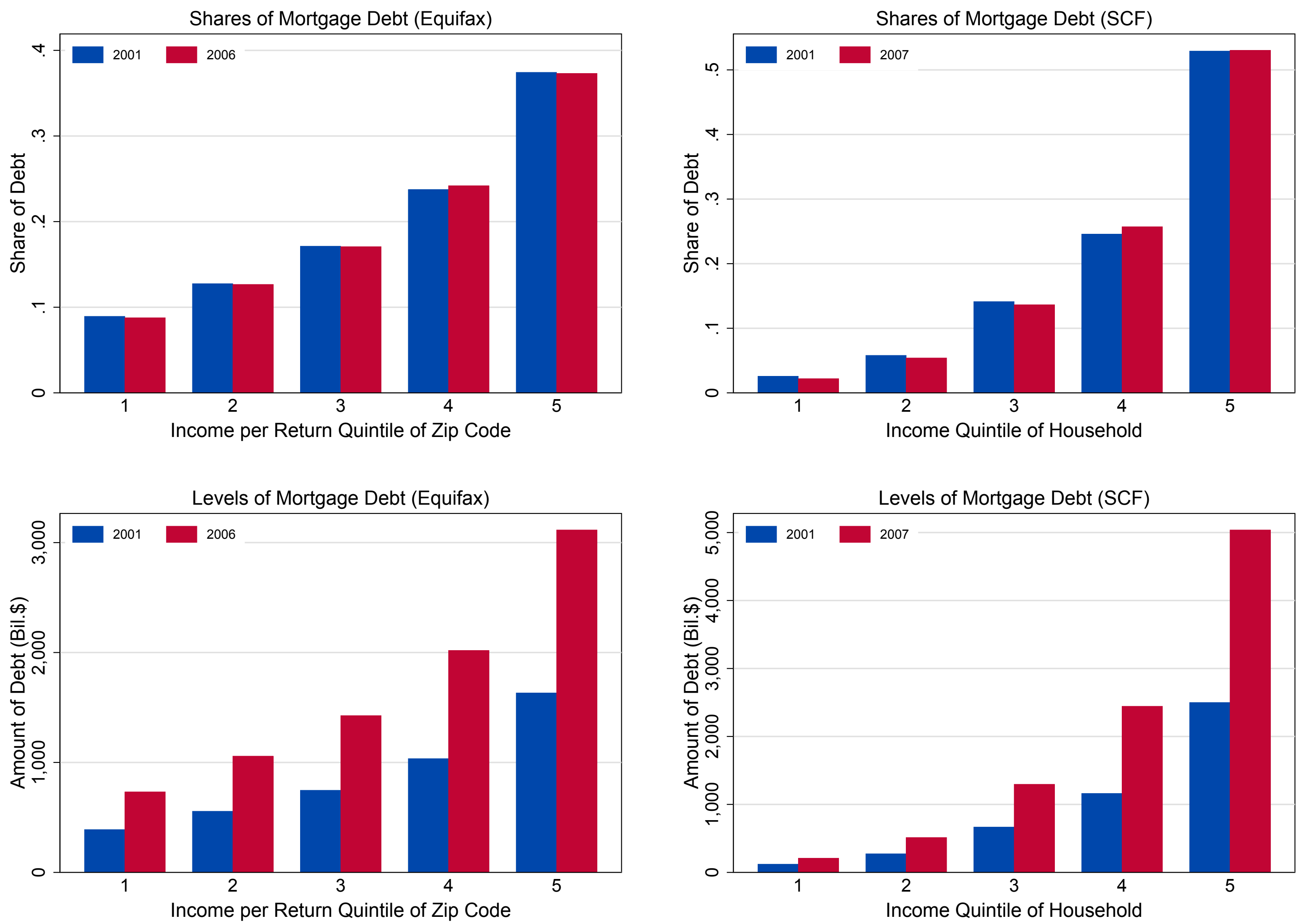


Figure A.4. Distributions of Mortgage Debt With Respect to Adjusted Gross Income (for Zip Codes) and Total Income (for Households). *Note:* The income measure used throughout the main text is salary and wage income. This figure uses AGI as the income measure for zip codes in the left panels, and total income from the SCF for households in the right panels. *Source:* NY Fed Consumer Credit Panel/Equifax, IRS Statistics of Income, and Survey of Consumer Finances.

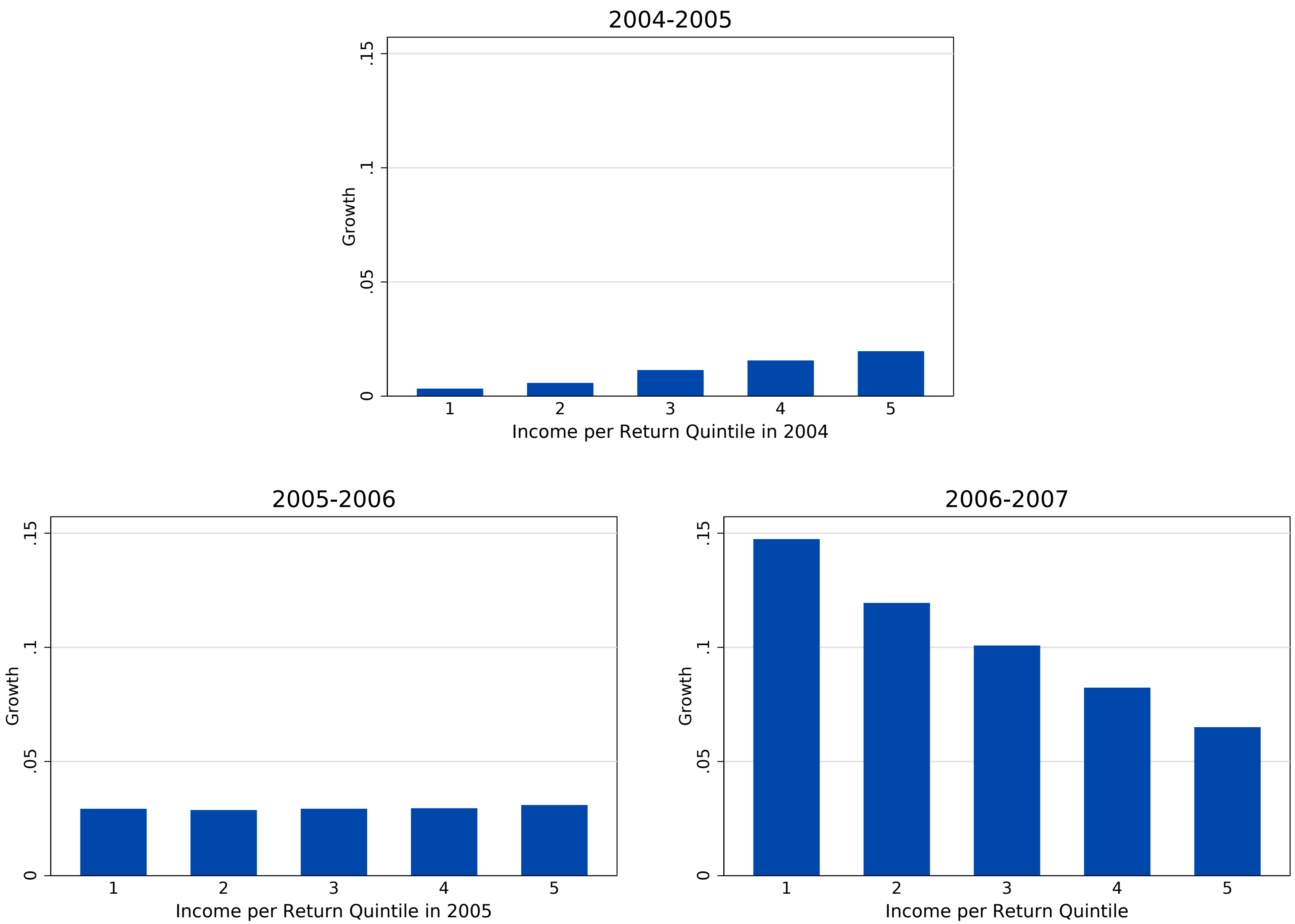


Figure A.5. Growth in IRS Returns by Income Quintile. *Note:* The top panel shows the growth in the total number of zip code-level tax returns between 2004 and 2005, grouped by zip code-level income in 2004. The bottom panels provide analogous information for returns growth in 2005–2006 and 2006–2007. The bottom right panel shows the strong inverse relationship between zip code-level returns growth and income between 2006 and 2007 that was generated by a surge of low-income persons who filed solely to take advantage of the 2007 tax stimulus. *Source:* IRS Statistics of Income.

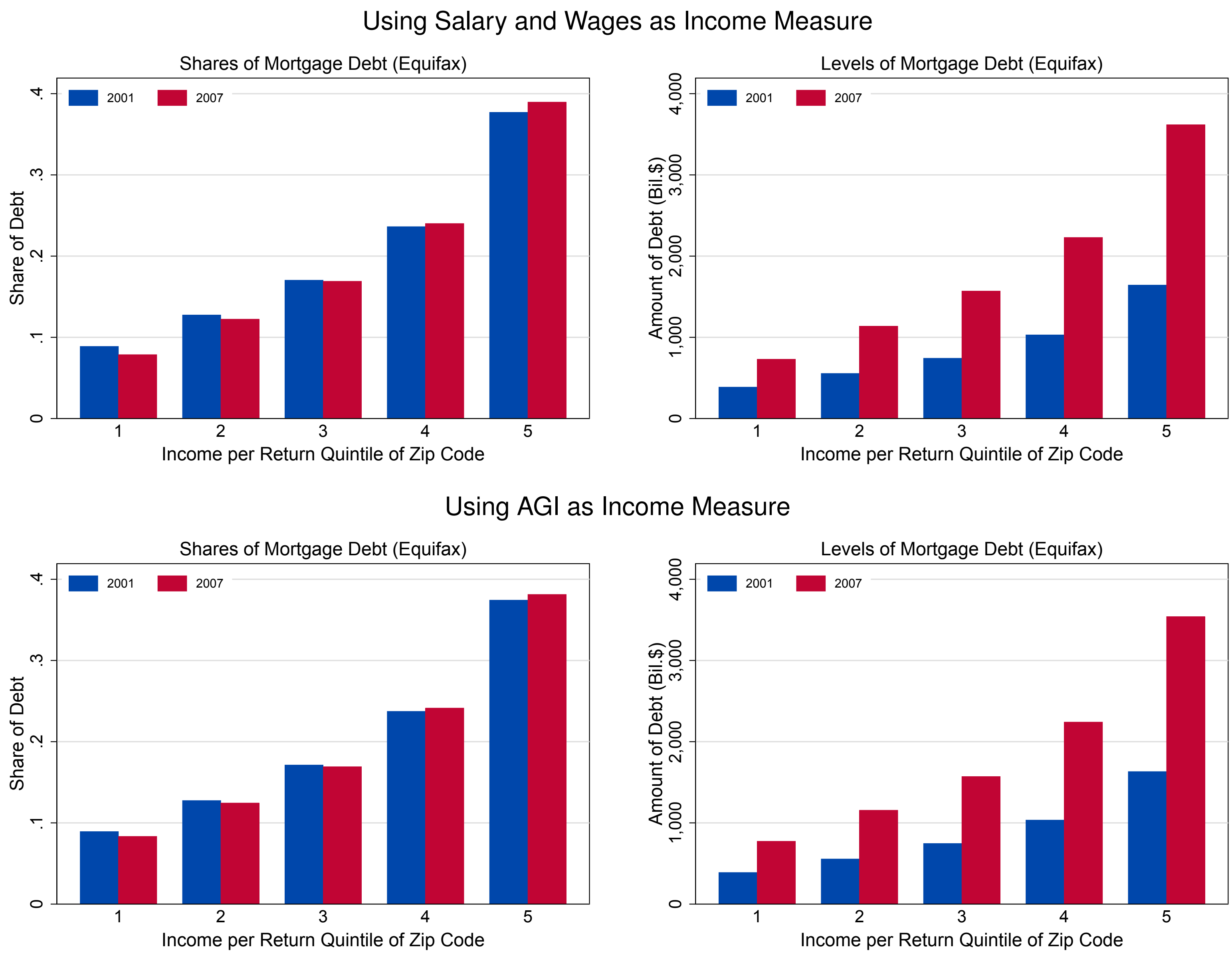


Figure A.6. Equifax/IRS Distributions of Debt for 2001 and 2007 using Alternative Income Definitions. *Note:* These graphs are analogous to the Equifax/IRS zip code-level bar charts in Figure 2, which depict distributions for 2001 and 2006, rather than 2001 and 2007. The lower panels in this figure also use AGI rather than wage and salary income. *Source:* NY Fed Consumer Credit Panel/Equifax, IRS Statistics of Income, and Survey of Consumer Finances.

Figure A.7. Distributions of Mortgage Debt and Real Estate Assets across 20 Categories of Salary and Wage Income in the SCF. *Note:* The two panels at left replicate the bar charts that use SCF data in Figure 2, but they divide households into 20 bins of wage and salary income rather than five. The two panels on the right plot the levels and shares of real estate assets held by each group. *Source:* Survey of Consumer Finances.

Poterba & Sinai (2008)		Authors' Calculation		
Household Income Category		Zip-Code Income Quintile	AGI	SW
$< \$40,000$	23.4	1	47.3	45.3
\$40–\$75,000	66.1	2	55.4	55.0
\$75–\$125,000	85.5	3	63.3	63.3
\$125–\$250,000	98.4	4	69.6	69.7
$\geq \$250,000$	99.9	5	72.8	73.7
All	63.1	All	63.8	63.8

Table A.1. Percentage of Homeowners and Mortgagors that Itemize Deductions by Income Categories. *Note:* The first two columns are from Table 1 of Poterba and Sinai (2008). These columns report the fractions of homeowners in the 2004 SCF who are predicted to have itemized their deductions, across various categories of income. These predictions are generated by comparing the value of the standard deduction to value of itemized deductions, where the latter is calculated for each homeowner in the 2004 SCF using the NBER's TAXSIM model. The remaining columns report the authors' calculations of zip code-level fractions of mortgaged households in 2007 who deducted their mortgage interest payments. Income quintiles are based on AGI in column (4) and salary and wage income in column (5). *Source:* Poterba and Sinai (2008), NY Fed Consumer Credit Panel/Equifax and IRS Statistics of Income.

Row		Bottom 95%		Top 5%	
		2001	2007	2001	2007
(1)	Share of Households	.95	.95	.05	.05
(2)	Share of Mortgage Debt	.78	.79	.23	.21
(3)	Average Mortgage Debt per Household ($)	36,348	68,270	201,178	348,084
(4)	Share of Total Dollar-Value Change in Debt: 2001–2007	.80		.20	
(5)	Average Income per Household ($)	47,317	56,062	492,450	628,213
(6)	Share of Total Household Income	.65	.63	.35	.37
(7)	Average Debt-to-Income Ratio (DTI)	.77	1.22	.41	.55
(8)	Absolute Change in DTI: 2001–2007	.45		.14	
(9)	Ln Change in DTI	.46		.30	
(10)	Ln Change in Mortgage Debt per Household	.63		.55	
(11)	Ln Change in Income per Household	.17		.24	

Table A.2. Mortgage Debt and Total Household Income by Income Category in 2001 and 2007. Note: To facilitate comparisons with Kumhof, Rancière, and Winant (2015), income is defined as total income for the household, not salary and wages. Households with no income are omitted. *Source:* Survey of Consumer Finances.

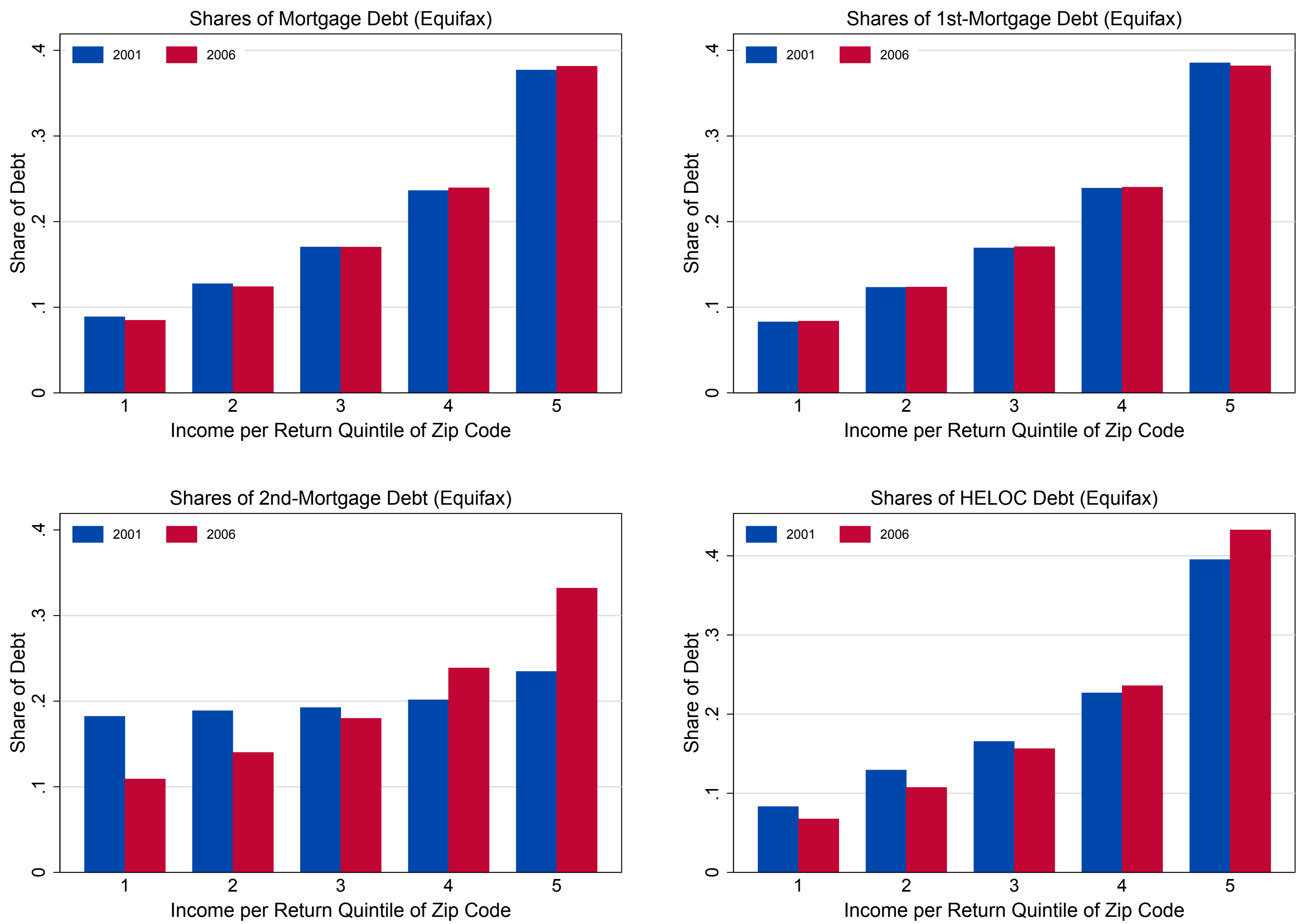


Figure A.8. Equifax/IRS Distributions of Debt by Mortgage Type, using Salary and Wages as Income Definition. *Note:* First mortgages include all purchase and refinance mortgages that are neither home equity loans nor home equity lines of credit (HELOCs). *Source:* NY Fed Consumer Credit Panel/Equifax and IRS Statistics of Income.

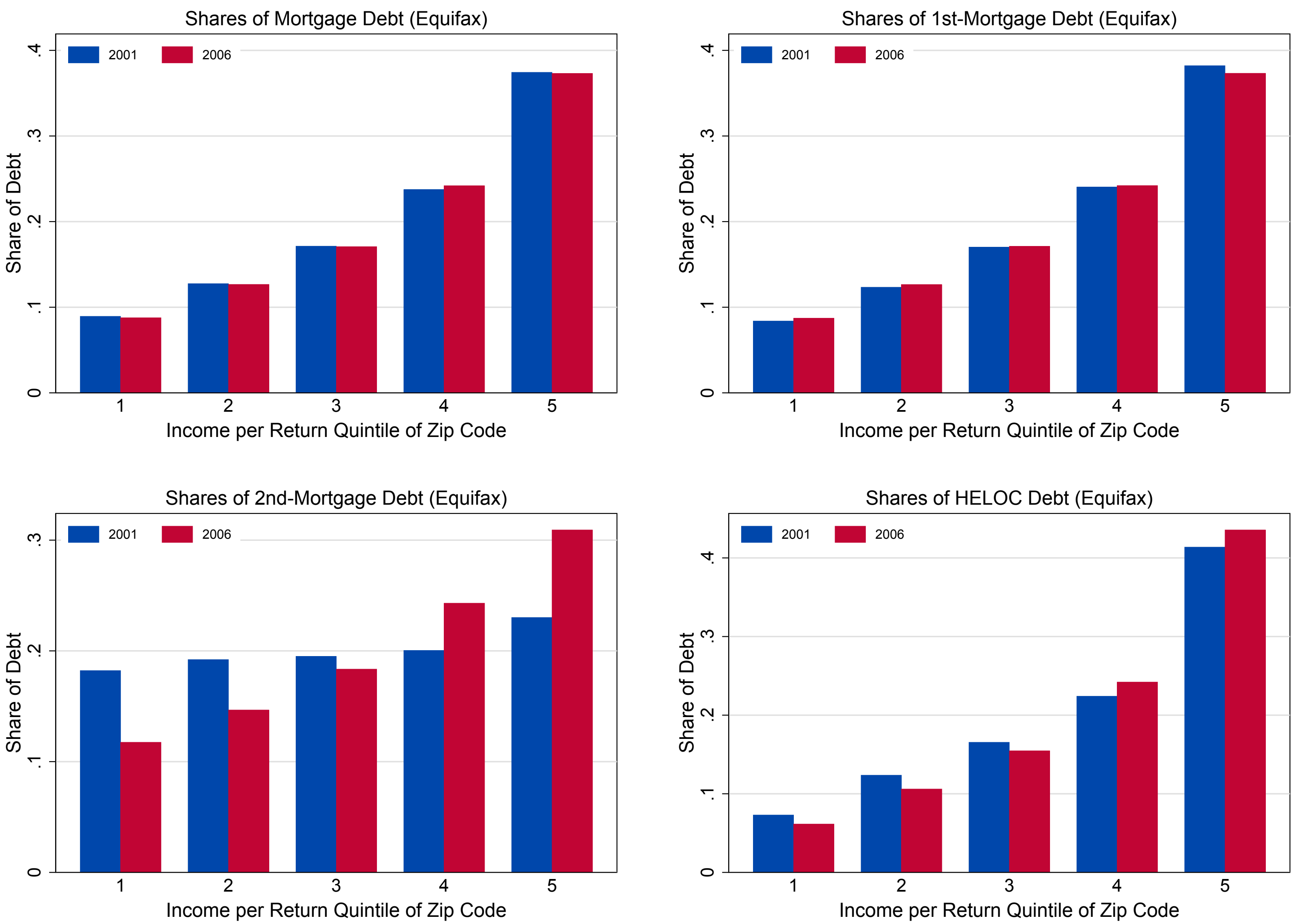


Figure A.9. Equifax/IRS Distributions of Debt by Mortgage Type, using AGI as Income Definition. *Note:* First mortgages include all purchase and refinance mortgages that are neither home equity loans nor home equity lines of credit (HELOCs). *Source:* NY Fed Consumer Credit Panel/Equifax and IRS Statistics of Income.

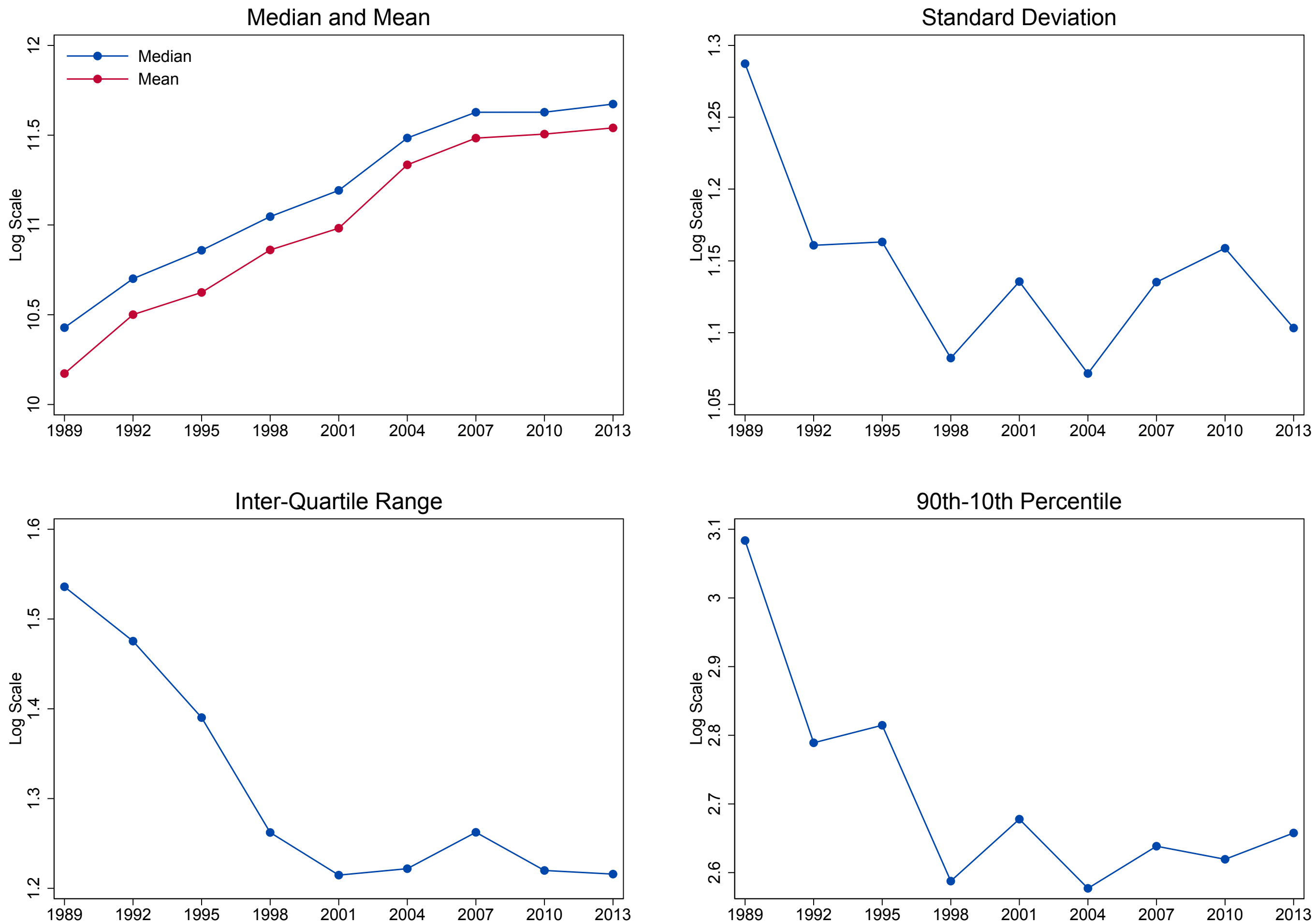
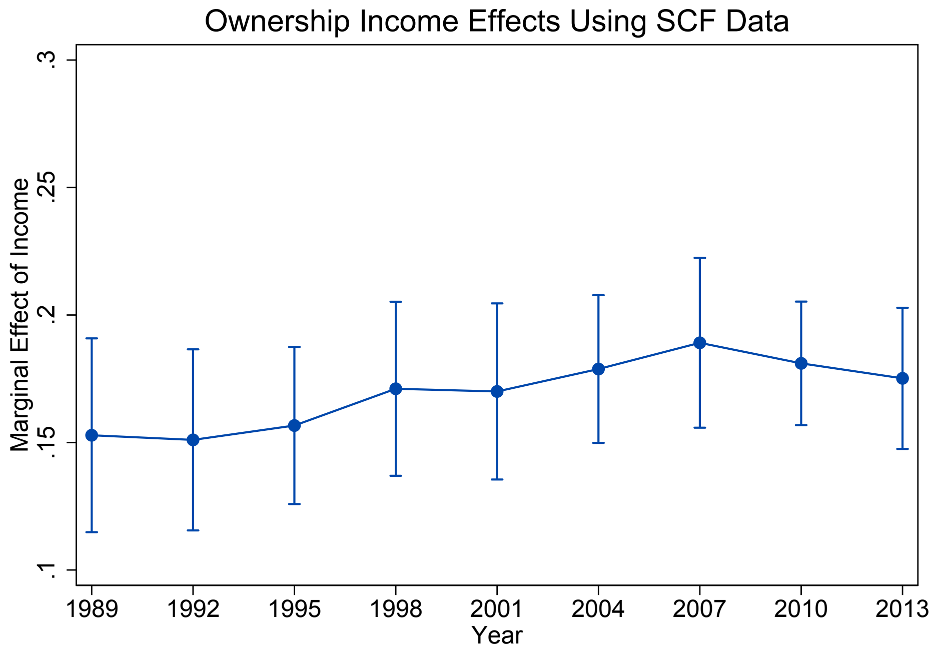


Figure A.10. Distributional Statistics for Household-Level Mortgage Debt in the Survey of Consumer Finances. *Note:* Statistics in these panels relate to the central tendency and dispersion in the distributions of household-level mortgage debt as measured in the Survey of Consumer Finances. Kernel estimates of these distributions for 1995, 2001, and 2007 appear in the top left panel of Figure 4. *Source:* Survey of Consumer Finances.



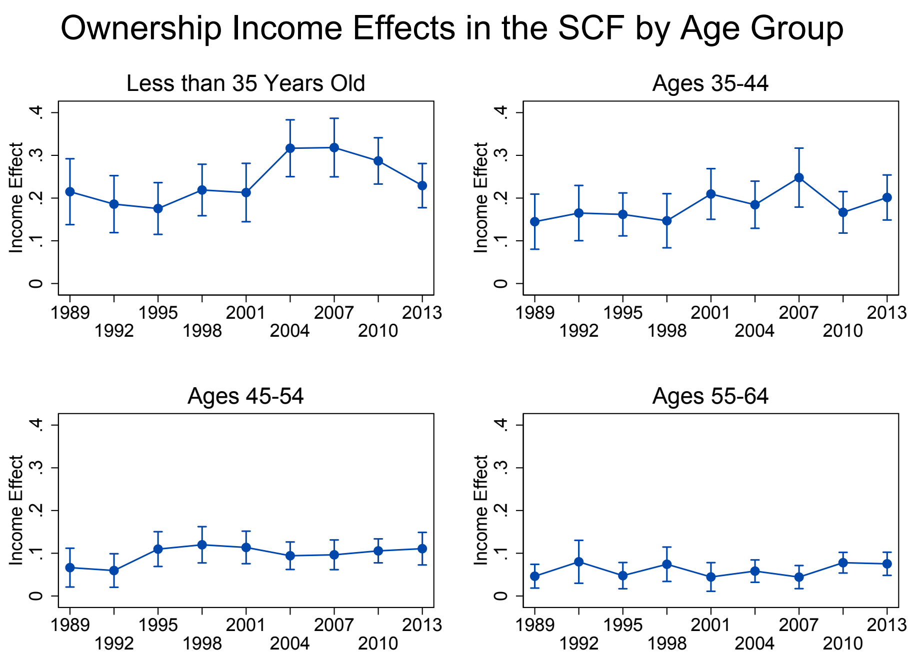


Figure A.11. Regression Evidence on the Relationship between Homeownership and Income in the Survey of Consumer Finances. *Note:* These panels are derived from logit homeownership regressions with the same right-hand-side variables and sample restrictions as those used for the Poisson regressions for total debt depicted in the bottom panel of Figure 5 and the SCF mortgageship regressions in Figure 7. All panels plot the marginal effect of log wage income on the probability of homeownership for all individuals (top panel) or for specific age groups (lower panels). The lower panels are generated from a regression in which the income regressor is interacted with dummy variables for the household head's age group. All income effects are marginal impacts on the probability of homeownership (not raw logit coefficients) and are calculated at the means of regressors from the first SCF implicate. *Source:* Survey of Consumer Finances.

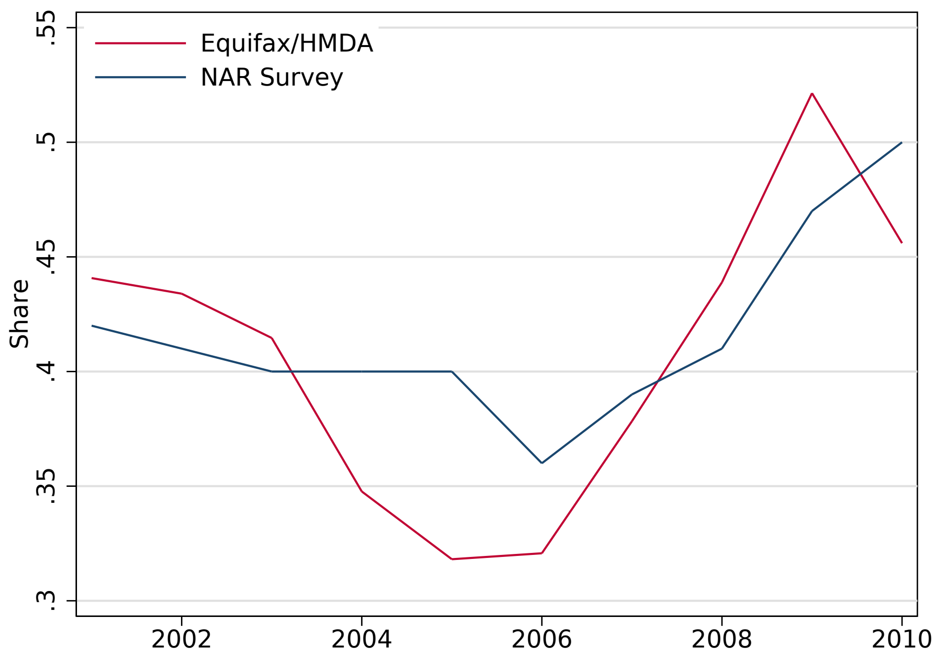


Figure A.12. Comparison of First-Time Borrower Share using Equifax and HMDA Data with First-Time Homebuyer Share Reported in the National Association of Realtors Survey. *Note:* The red line is the number of first-time mortgage borrowers in Equifax divided by the number of owner-occupied purchase mortgage originations from HMDA. The blue line is the share of homebuyers who are first-time homebuyers according to the National Association of Realtors Annual Survey. *Source:* NY Fed Consumer Credit Panel/Equifax and Internal, Home Mortgage Disclosure Act, and the National Association of Realtors.

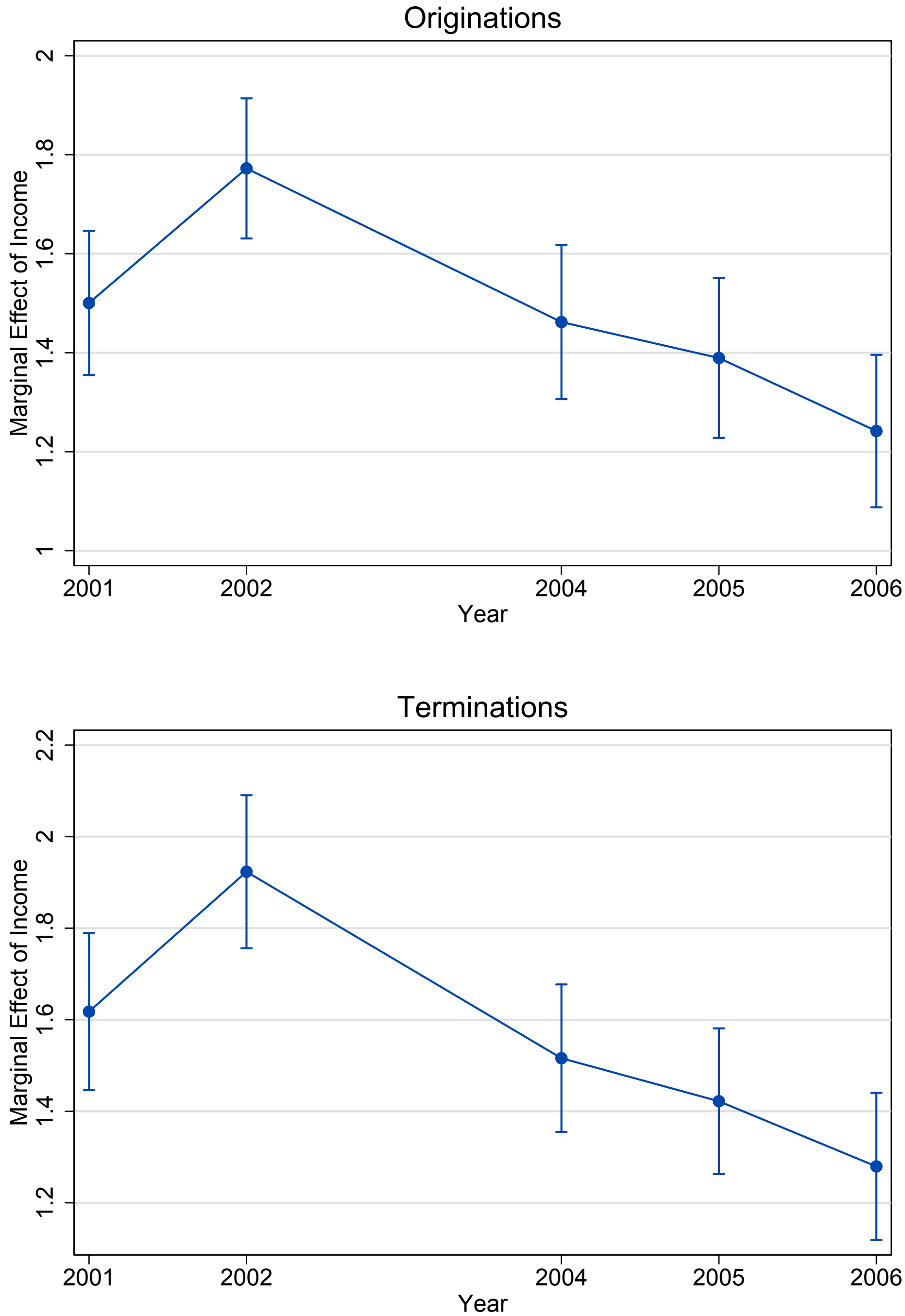
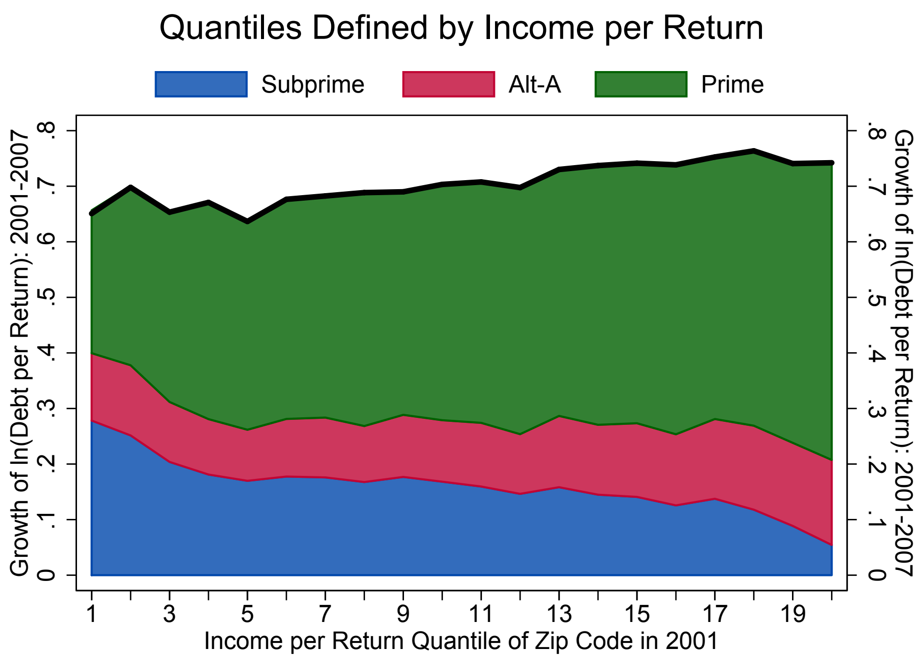


Figure A.13. Income Effects for Originations and Terminations without CBSA Fixed Effects. *Note:* Figure 10 in the main text displays income effects for originations and terminations when CBSA fixed effects are included. *Source:* NY Fed Consumer Credit Panel/Equifax and Internal Revenue Service Statistics of Income.



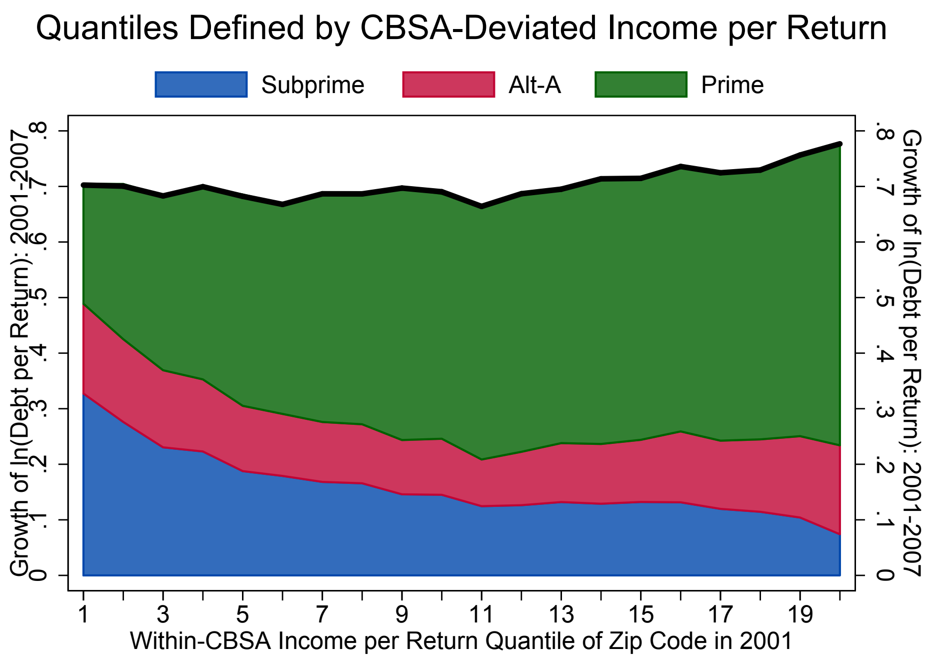


Figure A.14. Mortgage Debt Growth by Debt Type Across the Income Distribution of Zip Codes: 2001–2007. *Note:* These graphs are analogous to Figure 15, which is based on debt growth from 2001 to 2006 rather than growth from 2001 to 2007. *Source:* NY Fed Consumer Credit Panel/Equifax, CoreLogic Private Label Securities ABS Database, and IRS Statistics of Income.

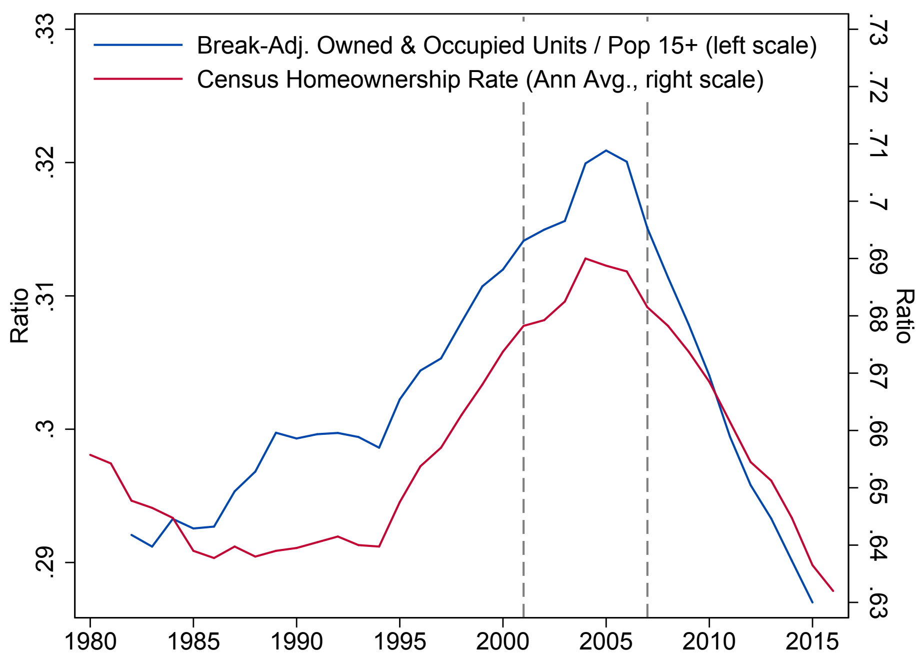


Figure A.15. Alternative Measures of Homeownwership. *Note:* The red line shows the standard homeownership rate, which is the share of occupied housing units that are occupied by owners. The blue line shows the total number of owner-occupied housing units divided by the population of adults aged 15 or older. The number of owner-occupied units is adjusted for break in 2000 using information on the total housing inventory in Table 953 of the 2004 Statistical Abstract of the United States. *Source:* Bureau of the Census and Bureau of Labor Statistics.