Appendix to Measuring Mortgage Credit Availability: A Frontier Estimation Approach

Elliot Anenberg Aurel Hizmo Edward Kung Raven Molloy*

May 15, 2019

A Online Appendix

A.1 Details of the HMDA to McDash/Corelogic Merge

The HMDA data are first restricted to first lien, purchase mortgages to be comparable with the McDash/CoreLogic sample.[1] Each HMDA loan is assigned a unique id ("hmdaid"). HMDA reports the census tract of the property whereas McDash/CoreLogic reports the zip code so the first step is to convert census tracts in HMDA into zip codes. We do this using the HUD-USPS Zip Crosswalk files and the Missouri Census Data center crosswalk for years in which the HUD-USPS Zip Crosswalk files are unavailable. This is a one-to-many merge, as census tracts can be contained in multiple zip codes, and so a single hmdaid may appear multiple times in the data after this initial merge.

Each McDash/CoreLogic loan is assigned a unique id ("mcdashid"). We then match mcdashid to all records in HMDA that have the same loan amount[2], the same zip code, and have origination dates within 45 days of each other. Flexibility on origination dates is permitted

*Anenberg, Hizmo and Molloy are at the Board of Governors of the Federal Reserve System; Kung is at UCLA. Emails: elliot.anenberg@frb.gov; aurel.hizmo@frb.gov; ekung@econ.ucla.edu; raven.s.molloy@frb.gov

[1]For the years 2001-2003, there is not a first lien flag. For these years, some junior liens are identified by finding loans that have the exact same borrower characteristics (income, sex, race, ethnicity), census tract, occupancy status, origination date, and selecting the loan origination where the loan amount is a small fraction of the larger loan amount.

[2]The loan amount in the McDash/CoreLogic data is first rounded to the nearest 1000 because all loan amounts in HMDA are rounded to the nearest 1000.

because some origination dates are missing in McDash/CoreLogic and must be imputed using the closing date of the loan. There could also be recording errors. In the case that a single hmdaid matches to more than one mcdashid, all potential matches for a particular hmdaid are sorted on difference in origination date, difference in occupancy status, and difference in loan type (e.g. FHA, GSE), in that order. Only the best potential match by this sort criteria is kept; the rest are dropped. This ensures that a single hmdaid does not match to more than one mcdashid. Then, in the case of where a mcdashid matches to more than one hmdaid, matches are again sorted on difference in origination date, difference in occupancy status, and difference in loan type, in that order. The first record in the sort is kept as a match.

In the case where a mcdashid does not match to any hmdaid, we then do a second round of matching that follows the same procedure as the above paragraph, except we permit zip codes to match on only the first 4 digits of the zip code. Flexibility in the match on zip code is permitted because some error is introduced when translating census tracts to zip codes. There could also be recording errors. All hmdaids and mcdashids that are matched in the first round are excluded from the second round. In the end, we matched 90 - 98 percent of all loans in the McDash/CoreLogic sample, depending on the year.

The next step is to collect all junior liens associated with each first lien mortgage origination at the time of origination. We follow the following procedure. For each first lien mortgage origination, we have all the borrower characteristics and property characteristics available in HMDA from the match described above. Therefore, we can match each first lien purchase origination with all junior lien purchase originations in HMDA that have the exact same census tract, origination date, occupancy status, and borrower characteristics (income, race, ethnicity, sex). A match between a first lien and junior lien where the junior lien loan amount is greater than the first lien loan amount, or where the combined LTV $>$ 120 is dropped. In practice, we find that there are very few instances where a single junior lien matches to multiple first lien originations. The share of originations that can be linked to a junior lien for the years 2001-2014 are reported in Table A5. To assess the quality of our lien matching algorithm, we compare the share of originations that can be linked to a junior lien by year in our data with the share computed using data from the Federal Reserve Bank of New York's (FRBNY) Consumer Credit Panel (CCP). Using these data, we compute individuals opening a small and large mortgage simultaneously and moving as a fraction of all who open a mortgage and move.

These shares by year are reported in the second column of Table A5. The shares in our data are similar to the ones coming from the FRBNY CCP data.

A.2 Confidence Intervals for the Aggregate Loan Frontiers

Figure A5 shows that the aggregate loan frontiers are fairly precisely estimated. The figure shows the estimated loan frontiers for various metro areas, along with 95% confidence intervals, which we computed using 100 bootstrapped repetitions. Confidence intervals are very tight, generally on the order of ±5% for the 100 largest metro areas that form our estimation sample. Beyond the 100th largest metro area, confidence intervals become larger, which reinforces our decision to restrict our analysis to the 100 largest metro areas.

A.3 Detail on Detecting Bunching at the Frontier Across Bins

For each of our fico, downpayment, income, year, msa bins that we compute frontiers for ("frontier bins"), we first calculate the share of observations within a certain distance of the frontier. We use twelve distance bins of length 5k, beginning at -49k (i.e. 44k-49k less than the frontier). Let s_{jb} denote the share of observations for frontier bin b within distance j of the frontier. Let j be the midpoint of the interval (e.g. for the interval [-4k,1k], $j = -1.5$). We then estimate the following regression:

$$s_{jb} = \alpha_0 + \alpha_1 j + \alpha_2 j^2 + \alpha_3 I[j = -1.5k] + \alpha_4 I[j > -1.5k] + \epsilon_{jb} \tag{1}$$

separately by group. $\alpha_3 > 0$ and $\alpha_4 < 0$ would be suggestive of bunching because it implies that the bin just before the frontier and the bins just after the frontier have more and less mass, respectively, relative to what a flexible function of j would suggest.[3] We define groups by first combining our 31 FICO frontier bins, 19 downpayment frontier bins, 18 income frontier bins, 14 year bins, and 100 frontier msas into 4 FICO bins (500-550, 550-600, etc), 6 downpayment bins (0-50k, 50k-100k,etc), 5 income bins (0-70k, 70k-110k, 110k-150k, etc), 14 year bins (i.e. years are not further grouped) and 10 city bins (cities are divided into bins according to their population rank). Each unique fico/downpayment/income/city/year bin combination consti-

[3] We also tried including higher order j terms, and the results were very similar.

tutes a group, so we have 16800 groups (4*6*5*14*10). We find that 75 percent of groups have α_3 statistically significantly greater than zero and α_4 statistically significantly less than zero at the ten percent level, indicating that bunching is fairly widespread across frontier bins.

A.4 Identification and Model Fit for Parametric Model

The likelihood function we maximize in the parametric model in Section 5.3 is

$$L = \prod_t \prod_i f_c(c_{ijt}) + f_d(d_{ijt}) \tag{2}$$

where

$$f_c(c_{ijt}) = \frac{1}{\sigma_{c,t}} \phi \left(\frac{c_{ijt} - E(c_{ijt})}{\sigma_{c,t}} \right) \Phi \left(\frac{\rho(c_{ijt} - E(c_{ijt}))}{\sigma_{c,t}\sqrt{1-\rho^2}} - \frac{d_{ijt} - E(d_{ijt})}{\sigma_{d,t}\sqrt{1-\rho^2}} \right) \tag{3}$$

and

$$f_d(d) = \frac{1}{\sigma_{d,t}} \phi \left(\frac{d_{ijt} - E(d_{ijt})}{\sigma_{d,t}} \right) \Phi \left(\frac{\rho(d_{ijt} - E(d_{ijt}))}{\sigma_{d,t}\sqrt{1-\rho^2}} - \frac{c_{ijt} - E(c_{ijt})}{\sigma_{c,t}\sqrt{1-\rho^2}} \right) \tag{4}$$

ϕ and Φ are the normal pdf and cdf, respectively.

To illustrate how the shape restrictions imposed in the parametric model described in Section 5.3 allow us to separately identify the demand and constraint distributions, consider Figure A10, which illustrates the distributions of unconstrained demand, borrowing constraints, and observed mortgage originations.[4] The figure shows that the distribution of originated loan amounts closely follows the distribution of unconstrained demand on the left tail, while it follows the distribution of constraints more closely on the right tail. Intuitively, borrowers with small loan amounts are unlikely to be constrained (conditional on characteristics), and the distribution of small loans more closely reflects the distribution of unconstrained demand. Borrowers with large loan amounts are more likely to be constrained, and the distribution of large loan amounts will more closely reflect the distribution of constraints. Figure A10 also illustrates the intuition for how we may identify changes to constraints over time separately from changes to demand. Under the distributional assumptions in the previous paragraph, if the left tail of the loan distribution remains the same from one period to the next, while the right tail of the distribution changes, we can reasonably attribute these changes to changes in

[4]The means and variances of borrowing constraints and unconstrained demand are set so that the distribution of mortgage originations is close to the observed distribution of loans in Chicago in 2003.

the distribution of constraints.

In total we estimate about 1500 coefficients including metro by year fixed effects. Table A2 shows the main coefficients of interest. First we focus on the coefficients of the constraint equation. As expected, the coefficient on FICO for the constraint equation is positive. The coefficient is low during the boom years and almost doubles in size in the post financial crisis period, illustrating the fact that credit score became more important in determining mortgage credit supply post-crisis. The coefficient on down-payment is very close to zero for all the years likely due to the existence of zero, or close to zero down-payment programs throughout our sample (also metro by year fixed effects soak up a lot of the variation in loan amounts and down-payments across locations). The coefficient on income is positive and relatively stable throughout the sample.

Turning to the coefficients of the demand equation, it is not as clear how to interpret the coefficients on FICO. We don't have a strong prior on what this coefficient should be or how it should evolve since it is not clear how demand for a loan is related to credit scores. The coefficients on down-payment are small and positive although larger than the coefficients on the constraint equation. The coefficients on income are positive over all the years and increase in the post crisis era indicating that demand for loans has increased more for higher income people in the post crisis era. The coefficients on income are smaller in the demand equation than the constraint equation indicating that income is a more important factor in determining credit supply than credit demand.

With regard to the model fit of the parametric model at our paramater estimates presented in Section 5.3, Figure A11 shows that the estimated model does a good job of fitting the empirical distributions of loan amounts in each year. We also verified the model fit for the distribution of loan amounts at the MSA level.

The parametric model also produces an estimate of the the share of borrowers who are bound by their constraints in each year; that is, the share of borrowers for whom $d_i \geq c_i$. This is a feature of the data that we do not explicitly target in estimation. Figure A12 shows this predicted share by year. As expected, the share is negatively correlated with the frontier, suggesting that when lending constraints are looser, the share of constrained borrowers is lower. The model predicts that 60 to 70 percent of borrowers take out the maximum obtainable loan amount given their FICO, income, and downpayment. We are not aware of any rigorous attempts to measure the share of constrained borrowers, but our estimates are similar in magnitude to other indirect and ad-hoc measures. Using the Survey of Consumer Expectations, Fuster and Zafar (2015) show that 42% of respondents would increase their demanded house value if downpayment requirements decreased from 20% to 5%. Applying institutional mortgage rules to the NLSY, Barakova et al. (2014) estimate that 58% of homeowners in 2003 and 72% in 2007 borrowed the maximum amount allowable.

A.5 Robustness Results for Section 6

We show that our estimates are both qualitatively and quantitatively robust to (i) alternative choices of m when computing the frontier, (ii) alternative choices of weights s_j^k in computing the instrument, (iii) using only full-documentation loans to reduce the bias associated with income misreporting, and (iv) computing the frontier conditioning on unobserved borrower heterogeneity, defined as the residual from an interest rate regression.

First, we test the robustness of our main results to our choice of m, which as explained in the text, is the number of draws one takes from the sample when computing the expected maximum loan amount. Table A4 shows results for $m = 500$ and $m = 2,000$. The results do not appear to be sensitive to our choice of m.

Second, we test the robustness of our main result to the choice of weights, s_j^k, used to compute the instrument. Column 3 of Table A4 shows the regression results when s_j^k is defined as the share of individuals in bin k in metro j in 2001, rather than averaged across time periods in our data. By fixing the weights using the data at the beginning of our sample period, we address potential concerns regarding households sorting over our sample period in a way that is affected by credit availability or housing market outcomes. The estimated elasticities of house price growth with respect to the frontier are comparable to those in the baseline specification.

Third, we re-estimate the frontier, dropping all loan originations that are not flagged as fully documented.[5] The motivation for this specification is that researchers have found that reported incomes in HMDA appear to be overstated, particularly in 2005 and 2006 (e.g. Avery et al. (2012), Blackburn and Vermilyea (2012)). By focusing on loans with full documentation, we are focusing on a sample for which income overstatement is less likely. Column 4 of Table

[5] In our data, 41% of loan originations are classified as fully documented, 15% are limited/no documentation, and 44% are of unknown documentation.

A4 shows that our results are similar when using this subsample of the data. We also verified that the general patterns of how the loan frontier expanded and contracted that we document in the main text are similar when we use this subsample of the data.

Finally, we consider the possibility of omitted variables. As discussed in Section 5, unobserved heterogeneity may be a concern if changes to the frontier are not correlated with changes to borrowing constraints faced by typical borrowers. In the IV regression, our instrument will be valid only if metro-by-year specific shocks to the distribution of unobservables (that also independently affect house prices) are not correlated across metro areas.[6] To address this concern, we construct the frontier using the borrower's residualized interest rate at the time of origination as an additional input.[7] The motivation for this approach is that one might expect that, conditional on observable characteristics, lower interest rates are available to borrowers with better unobserved characteristics. Then, the interest rate residual can be used as a proxy for borrower unobserved characteristics. We find that the frontier tends to be larger for metro/year/borrower type bins where the residual is more negative, which is consistent with this interpretation. To keep the analysis tractable, we categorize borrowers into two types: high types who have residual interest rates below average, and low types who have residual interest rates above average. Column 5 of Table A4 reports the results when we aggregate over the unobserved borrower type using equal weights for low and high types. The estimated elasticity of house price growth with respect to the frontier is comparable to the ones in our baseline specification, suggesting that changes in the distribution of borrower unobservables are not driving the estimation results.

A.6 Model of House Price Determination under Borrowing Constraints

We present here a simple model of house price deterimination in the presence of borrowing constraints. The purpose of the model is to provide a context for interpreting our estimates of the elasticity of house prices to the aggregate loan frontier presented in Section 6.

[6]Shocks to the distribution of unobservables that are correlated across metro areas would be captured by our fixed effects if the shocks are spread across all borrower types.

[7]In particular, we obtain the residual by regressing the interest rate at origination on FICO, LTV, income, origination amount, ARM dummy, loan type dummies, 30-year-term dummy, metro fixed effects, and interaction terms. The regressions are run separately for each year.

Consider a housing market where all housing units are identical and housing is indivisible, so that each consumer can purchase one and only unit of housing. Let the price of a unit be p. The housing supply curve is given by $S(p)$. There is a unit mass of potential buyers. Each buyer has a private valuation, v_i, for buying a house. If $v_i \geq p$ then the buyer will want to purchase, otherwise not. Buyers are also heterogeneous in the amount of downpayment they can make, d_i, and the maximum amount they are allowed to borrow, m_i. If $d_i + m_i \geq p$, then the buyer is able to make the purchase. Otherwise, the buyer is priced out of the market. In equilibrium, the mass of buyers making a purchase is equal to the supply of housing:

$$\int 1\left[v_i \geq p \text{ and } d_i + m_i \geq p\right] di = S(p)$$

To simplify the analysis, let us now assume that v_i is iid over support $[0, \infty)$ according to the cdf $F(v)$. Further assume that $d_i + m_i$ can be written as $\mu_d + \mu_m + \epsilon_i$, where:

$$\mu_d = E[d_i] \quad \mu_m = E[m_i] \quad E[\epsilon_i] = 0$$

and ϵ_i is iid over support $[-\mu_m - \mu_d, \infty)$ according to cdf $G(\epsilon)$. For simplicity, we assume that ϵ_i and v_i are independent. The equilibrium condition becomes:

$$\begin{aligned} \int_p^{+\infty} \int_{p-\mu_d-\mu_m}^{+\infty} g(\epsilon) f(v) \, d\epsilon dv &= S(p) \\ [1 - F(p)]\left[1 - G(p - \mu_d - \mu_m)\right] &= S(p) \end{aligned} \tag{5}$$

We are interested in how p varies with μ_m, the average amount of credit available to potential buyers. Taking first derivatives with respect to p and μ_m, we obtain:

$$\frac{dp}{d\mu_m} = \frac{[1 - F(p)]\, g(p - \mu_d - \mu_m)}{[1 - F(p)]\, g(p - \mu_d - \mu_m) + [1 - G(p - \mu_d - \mu_m)]\, f(p) + S'(p)}$$

A few observations to note:

1. $dp/d\mu_m$ will generally be less than 1, but will be close to 1 if supply is inelastic and v_i is large relative to $d_i + m_i$.

2. The price elasticity $\frac{dp}{d\mu_m}\frac{\mu_m}{p}$ can still be greater than 1 even if $dp/d\mu_m$ is less than 1. This

can happen if μ_m, the average borrowing constraint, is greater than p, which is true in our data.

3. $dp/d\mu_m$ is greater when housing supply is more inelastic. This is intuitive, and what we find in Table 3.

4. A high price elasticity does not require a significant fraction of borrowers to be actively constrained. In this model, the borrowers with binding constraints are the borrowers with $d_i + m_i = p$, which is an infinitessimal set. However, prices are responsive to μ_m despite not many borrowers being actively constrained, because it is the actively constrained borrowers who are marginal and therefore determining the equilibrium price.

Table A1: Growth Rate of Loan Amount Frontier for Selected Metropolitan Areas

	2001 to 2006	2006 to 2008	2008 to 2014
National Average	40.4	-19.3	-8.6
Atlanta	30.4	-16.4	-7.6
Boston	28.3	-16.4	-7.8
Chicago	32.8	-8.3	-22.6
Dallas	23.2	-13.7	-3.6
Denver	24.9	-16.8	-2.7
Detroit	22.0	-24.4	-3.8
Indianapolis	20.6	-15.8	-5.5
Las Vegas	61.7	-25.7	-13.9
Miami	67.9	-31.5	-6.9
New York	43.1	-22.7	-9.0
Phoenix	51.6	-24.6	-8.1
San Francisco	39.0	-31.4	-4.1
Seattle	42.3	-15.5	-9.3
Washington, DC	54.6	-22.2	-1.2

Table A2: The Parameter Estimates from the Parametric Model

Year	Coefficients of constraint equation				Coefficients of demand equation			
	$lnFICO$	$ln(1+DownPmt)$	$lnIncome$	σ_c	$lnFICO$	$ln(1+DownPmt)$	$lnIncome$	σ_d
2001	0.309	0.017	0.737	0.245	0.083	0.009	0.087	0.721
2002	0.291	0.012	0.760	0.250	-0.052	0.019	0.054	0.695
2003	0.283	0.009	0.742	0.252	-0.152	0.032	0.034	0.673
2004	0.268	-0.002	0.802	0.256	-0.029	0.031	0.024	0.654
2005	0.292	-0.012	0.829	0.248	0.078	0.005	0.032	0.638
2006	0.284	-0.011	0.821	0.240	0.373	-0.017	0.053	0.642
2007	0.313	-0.007	0.768	0.247	0.057	0.000	0.052	0.680
2008	0.453	0.008	0.768	0.250	-0.535	0.004	0.045	0.731
2009	0.469	0.006	0.785	0.256	-0.147	-0.006	0.079	0.720
2010	0.378	0.001	0.823	0.259	-0.117	-0.005	0.107	0.684
2011	0.450	-0.002	0.847	0.272	-0.546	0.017	0.109	0.709
2012	0.496	0.003	0.829	0.280	-0.832	0.031	0.108	0.707
2013	0.479	0.004	0.827	0.275	-0.801	0.029	0.105	0.685
2014	0.512	0.004	0.820	0.270	-0.584	0.025	0.122	0.675

Note: The model also includes CBSA by year fixed effects for both the constraint and demand equation, which we do not display here.

Table A3: First Stage Effects of the Instrument on Loan Frontiers

Dep. variable:	$\Delta lnFrontier$		$I_{Loose} \times \Delta lnFrontier$	$I_{Tight} \times \Delta lnFrontier$
	(1)	(2)	(3)	(4)
$\Delta lnInstrument$	0.657***	0.632***	-0.175***	0.058
	(0.100)	(0.121)	(0.055)	(0.097)
$I_{Loose} \times \Delta lnFrontier$		-0.070*	0.794***	0.078***
		(0.041)	(0.025)	(0.019)
$I_{Tight} \times \Delta lnFrontier$		0.069	0.001	1.097***
		(0.047)	(0.005)	(0.043)
ΔLn Delinquency Rate	-0.066***	-0.062***	-0.003	-0.030***
	(0.005)	(0.006)	(0.002)	(0.005)
ΔLn Income	0.214**	0.215**	0.025	0.116
	(0.095)	(0.089)	(0.023)	(0.084)
ΔLn Employment	0.425**	0.398**	0.046	0.306*
	(0.197)	(0.189)	(0.047)	(0.173)
Observations	1217	1152	1152	1152
R^2 overall	0.819	0.820	0.724	0.805
F statistic	416	354	82	184

Note: All the variables in this regression are in log differences. The sample consists of annual data from 2001 to 2013 for 100 metropolitan areas. The variable I_{Loose} indicates CBSAs in the top third of the housing supply elasticity and I_{Tight} indicates CBSAs in the bottom third of the housing supply elasticity. $Frontier$ is the loan frontier aggregated up to the metro-year level, as described in Section 4. $Instrument$ is a shift-share instrument for $Frontier$. All specifications include metro area and year fixed effects. Standard errors are clustered at the metro area level and are given in parentheses. *, **, *** indicate statistical significance at the 90%, 95%, and 99% level respectively.

Table A4: Robustness with respect to alternate specifications

	m=500	m=2000	Weights from 2001	Only Full Doc. Loans	Controlling for Unobs. Type
	(1)	(2)	(3)	(4)	(5)
$\Delta lnFrontier$	0.937***	1.057***	0.858*	0.933**	1.303***
	(0.221)	(0.268)	(0.470)	(0.461)	(0.357)
ΔLn Delinquency Rate	-0.096***	-0.091***	-0.109***	-0.111***	-0.074***
	(0.019)	(0.021)	(0.035)	(0.030)	(0.026)
ΔLn Income	-0.022	-0.061	0.022	0.001	-0.156
	(0.085)	(0.094)	(0.118)	(0.109)	(0.120)
ΔLn Employment	0.840***	0.818***	0.911***	0.713**	0.620***
	(0.206)	(0.208)	(0.221)	(0.291)	(0.226)
Observations	1217	1217	1217	1217	1217
R^2 overall	0.799	0.767	0.776	0.727	0.737

Note: All the variables in this regression are in log differences. The sample consists of annual data from 2001 to 2013 for 100 metropolitan areas. All specifications include metro area and year fixed effects. Standard errors are clustered at the metro area level and are given in parentheses. *, **, *** indicate statistical significance at the 90%, 95%, and 99% level respectively.

Table A5: Share of Originations Linked to a Junior Lien

Year	Our Data	FRBNY CCP/Equifax Data
2001	4.23	4.50
2002	5.78	6.03
2003	7.26	9.02
2004	13.66	13.48
2005	22.83	18.6
2006	26.99	22.41
2007	14.06	19.92
2008	2.36	6.63
2009	0.73	2.08
2010	0.85	1.65
2011	1.11	1.66
2012	1.05	1.75
2013	0.93	1.39
2014	1.42	1.87

Figure A1: Phoenix Loan Frontiers, 2004 and 2012

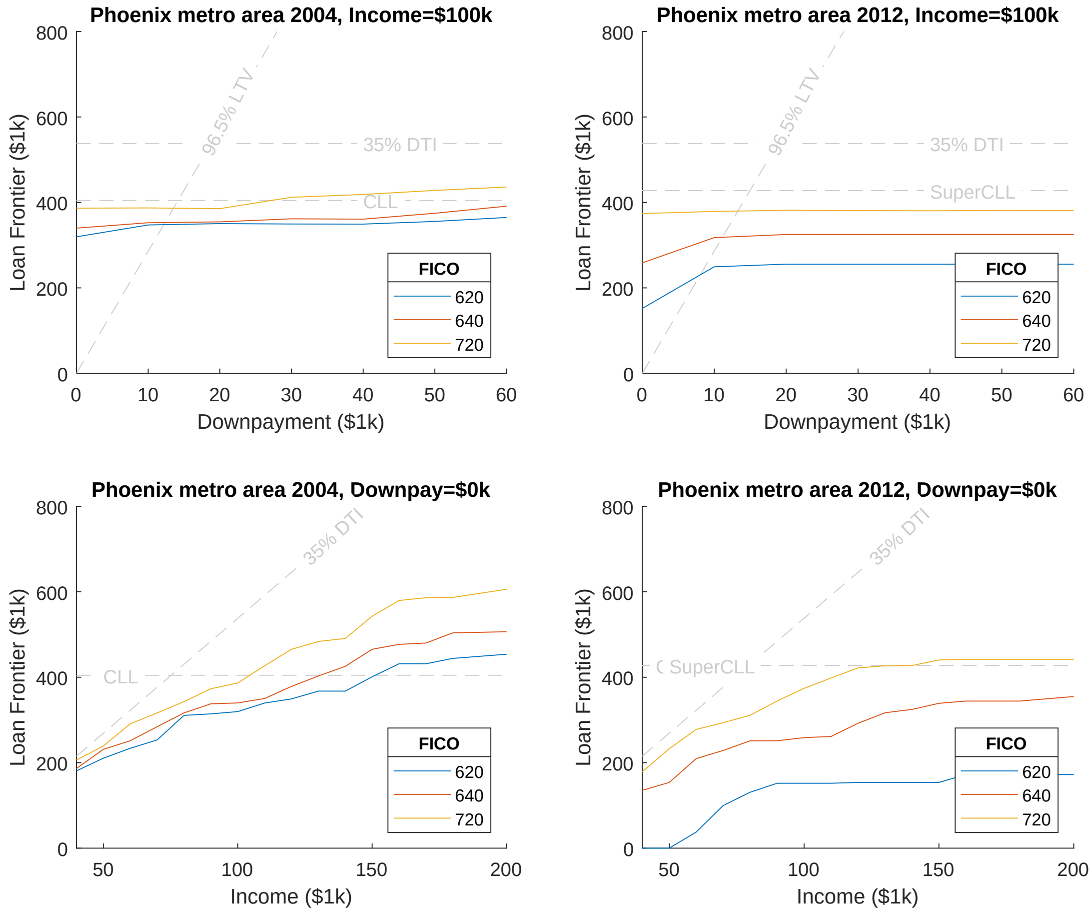


Figure A2: Washington DC Loan Frontiers, 2004 and 2012

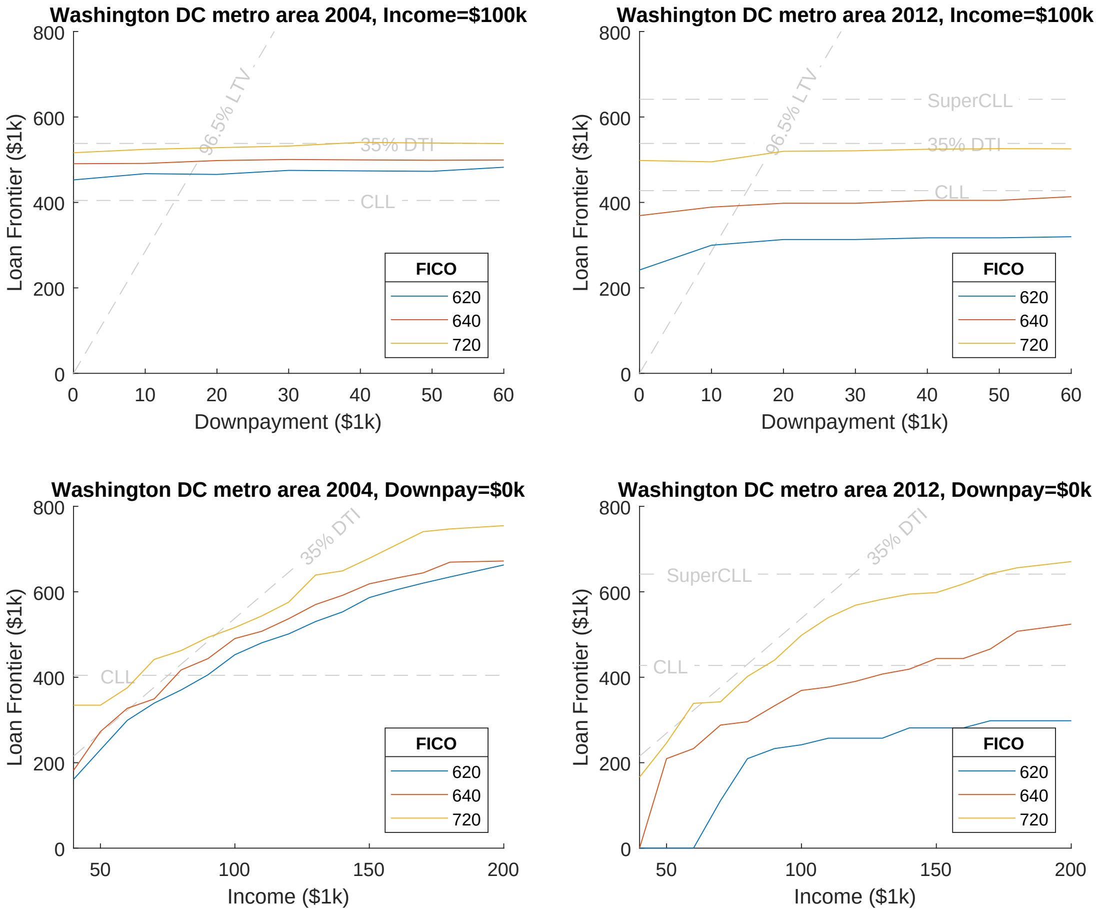


Figure A3: Chicago Loan Frontiers, 2004 and 2012

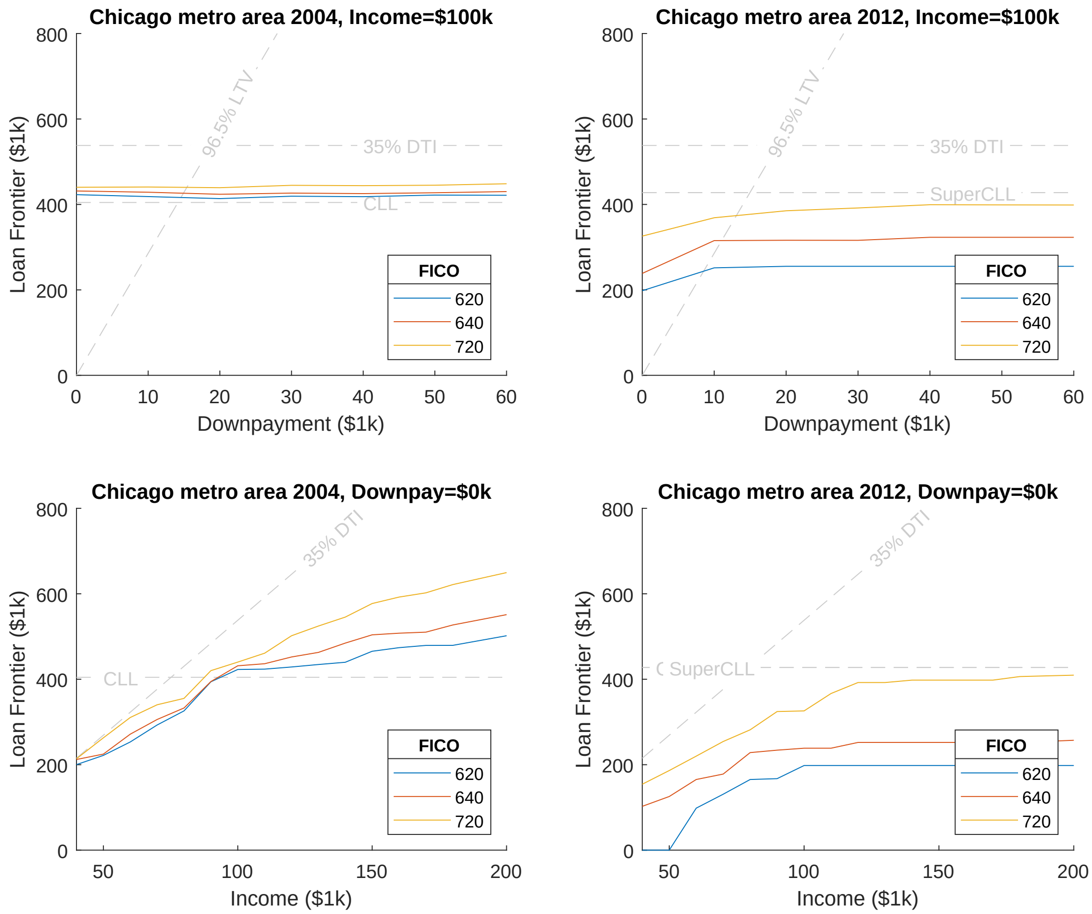


Figure A4: Boston Loan Frontiers, 2004 and 2012

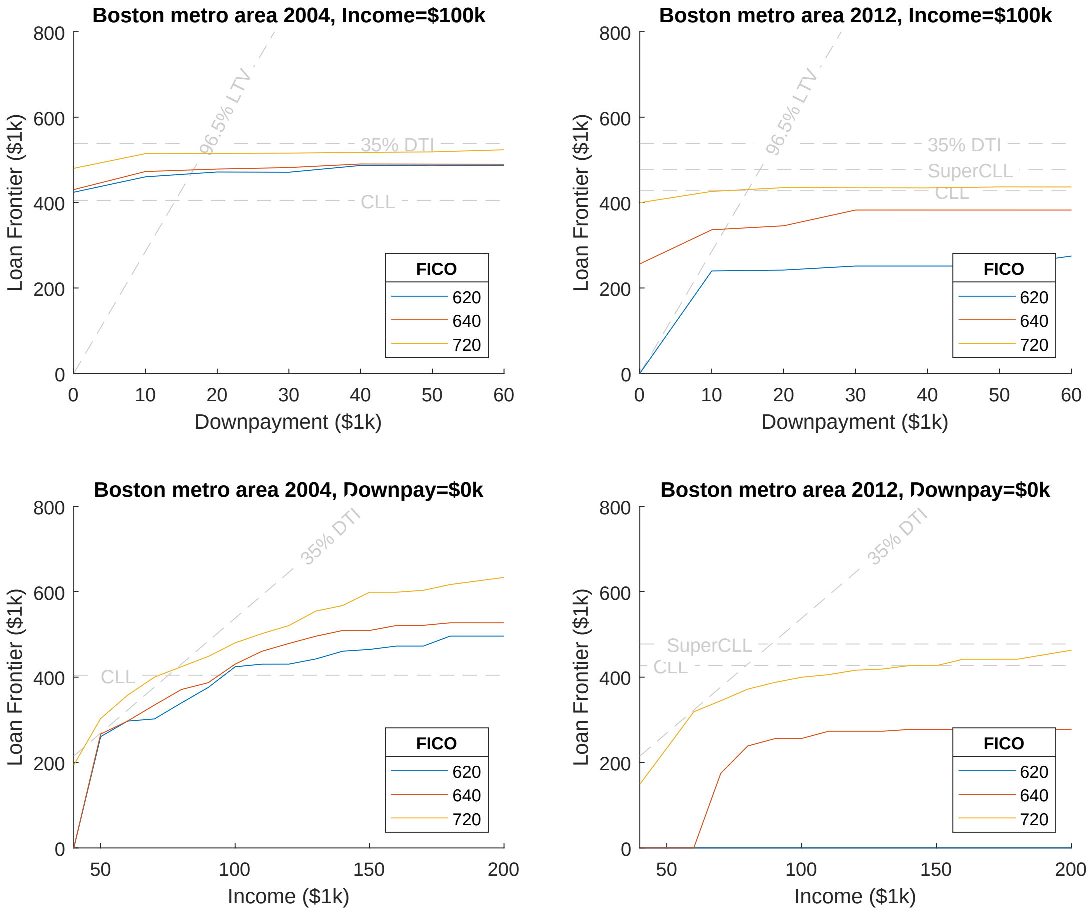
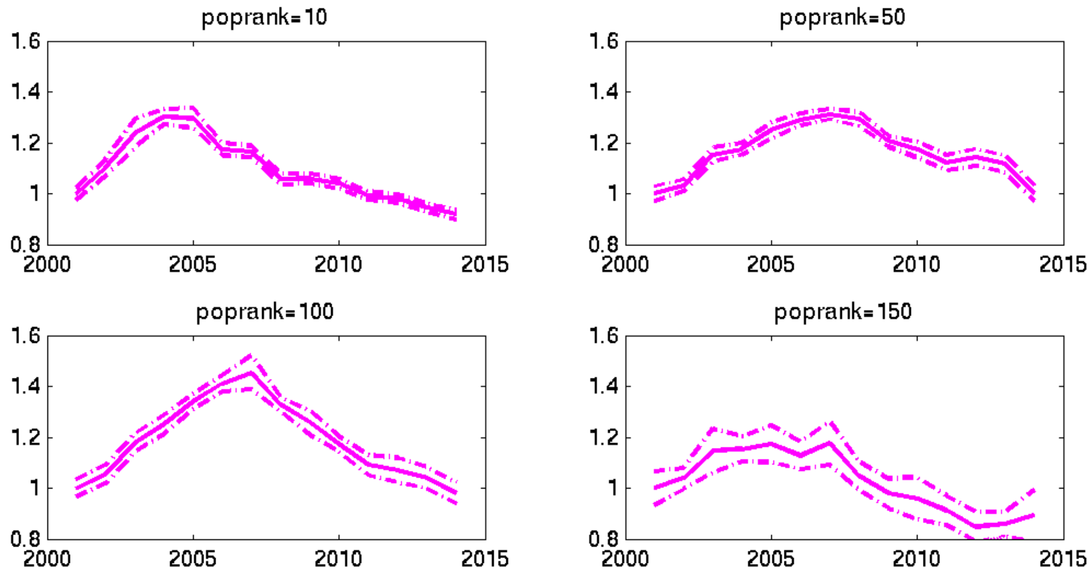
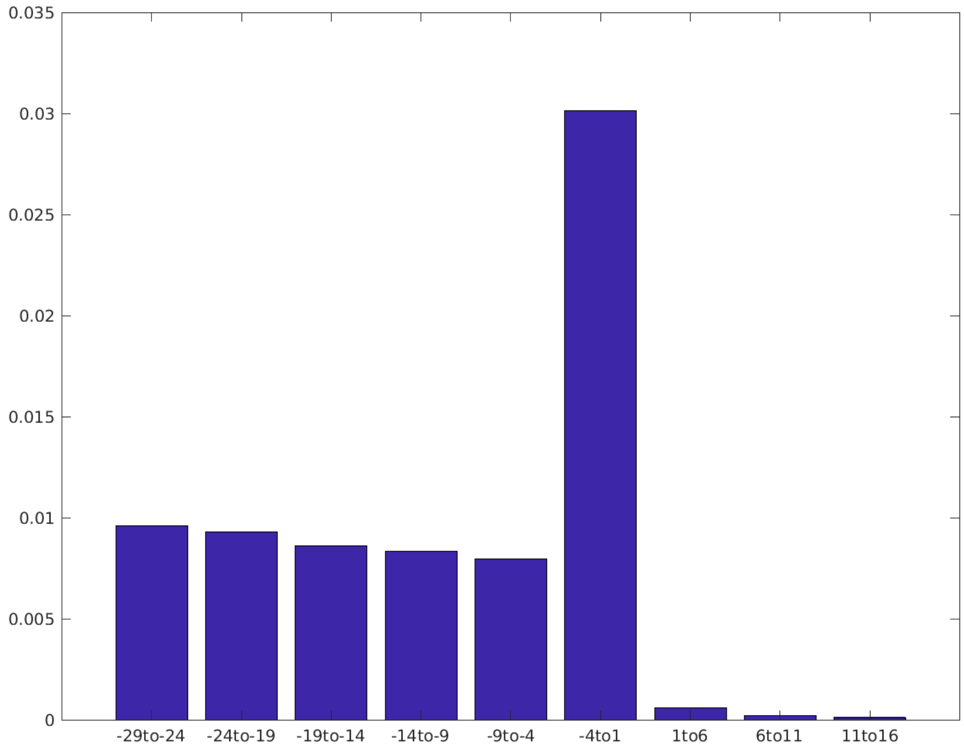


Figure A5: Confidence Intervals for Aggregate Loan Frontiers



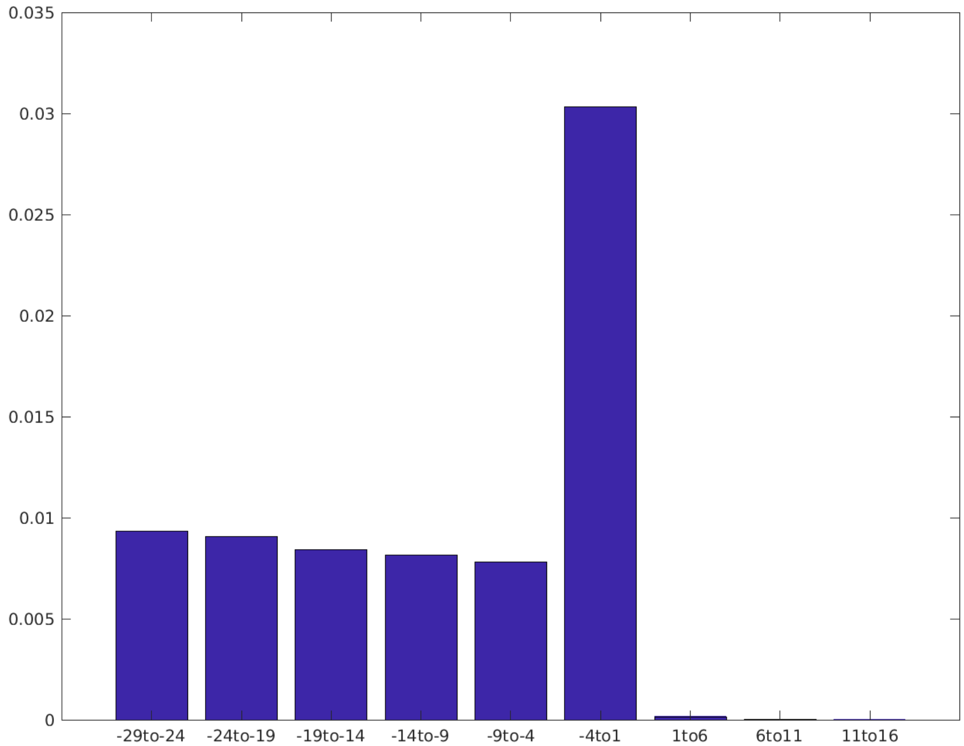
This figure shows 95 percent confidence intervals, in dotted lines, of the aggregate loan frontier, the solid line, for select MSAs. Poprank is the population rank of the MSA. The MSAs shown are Boston, MA; Salt Lake City, UT; Spokane, WA; South Bend, IN respectively.Confidence intervals are computed using 100 bootstrap repetitions. Loan frontiers are indexed to one in 2001.

Figure A6: Distribution of Mortgage Originations Around the Loan Frontier, $m = 500$



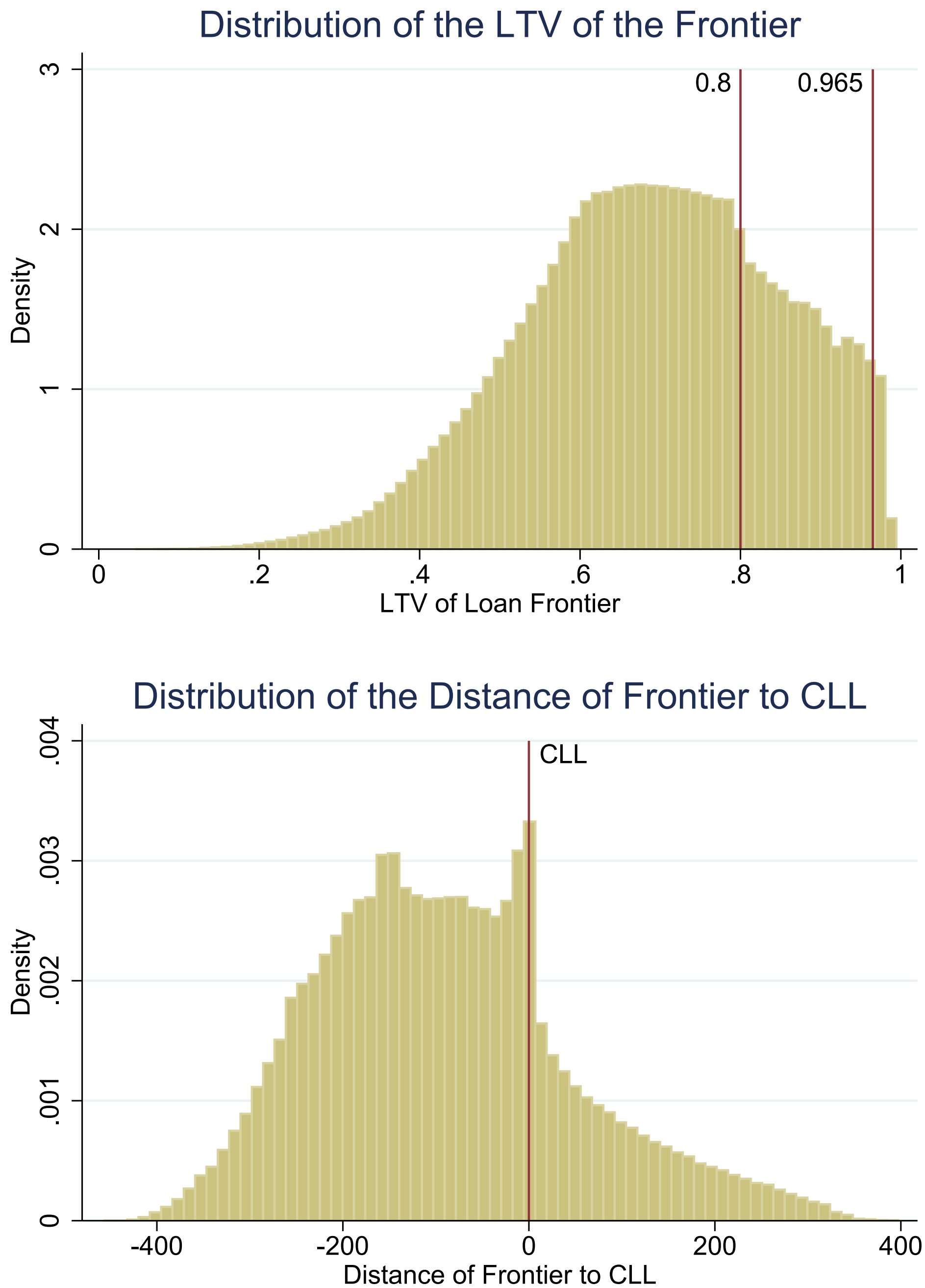
For each borrower type/year/metro area, we compute the share of observations within \$5,000 intervals around the estimated frontier for that borrower. The figure plots the histogram when we take the simple average of these shares across all borrower types, years, and metro areas.

Figure A7: Distribution of Mortgage Originations Around the Loan Frontier, $m = 2000$



For each borrower type/year/metro area, we compute the share of observations within \$5,000 intervals around the estimated frontier for that borrower. The figure plots the histogram when we take the simple average of these shares across all borrower types, years, and metro areas.

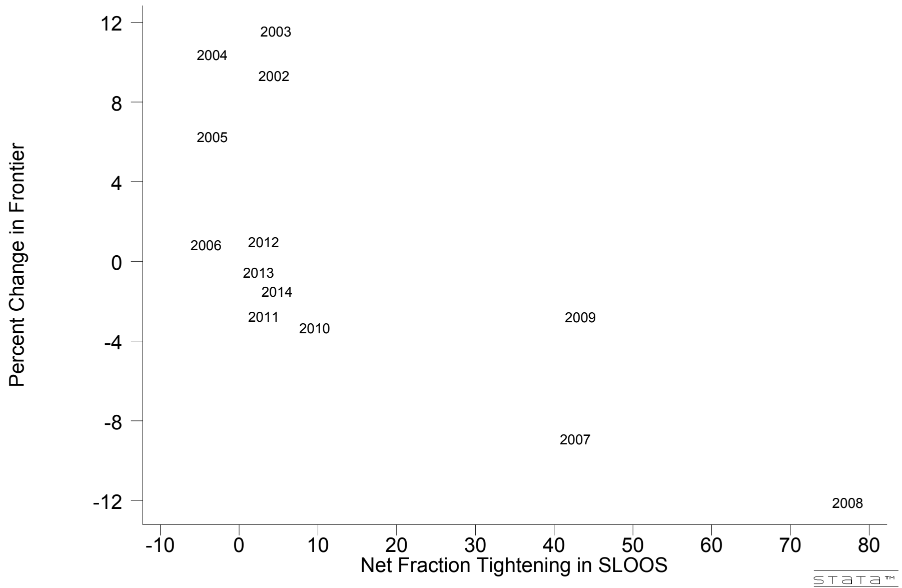
Figure A8: Distribution of the Frontier



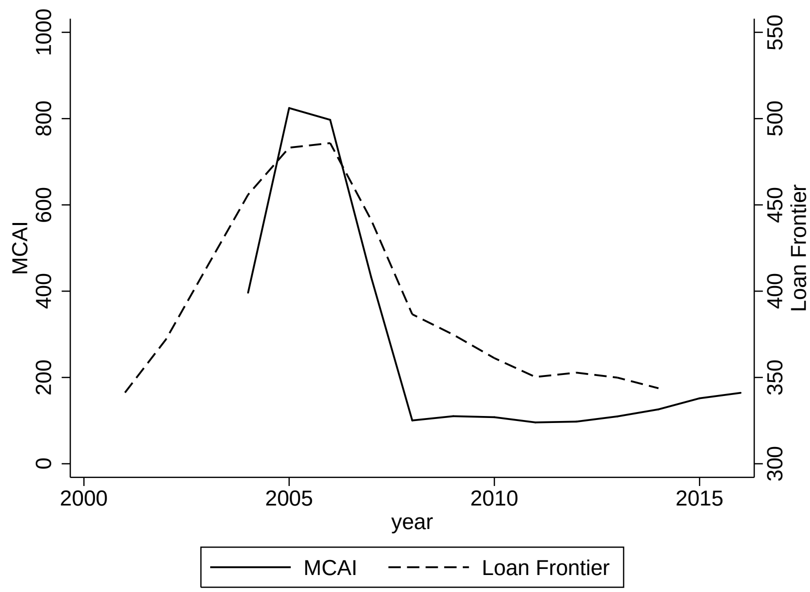
The distribution of the frontier across borrower types, metro areas, and years is plotted for two different ways of measuring the frontier, the implied LTV and the distance to conforming loan-limit. The LTV of the frontier is $LTV = \frac{frontier}{frontier+downpay}$ and the distance to CLL is simply the frontier minus the CLL.

Figure A9: Correlation Between the Loan Frontier and the SLOOS / MCAI

(a) Senior Loan Officer Opinion Survey

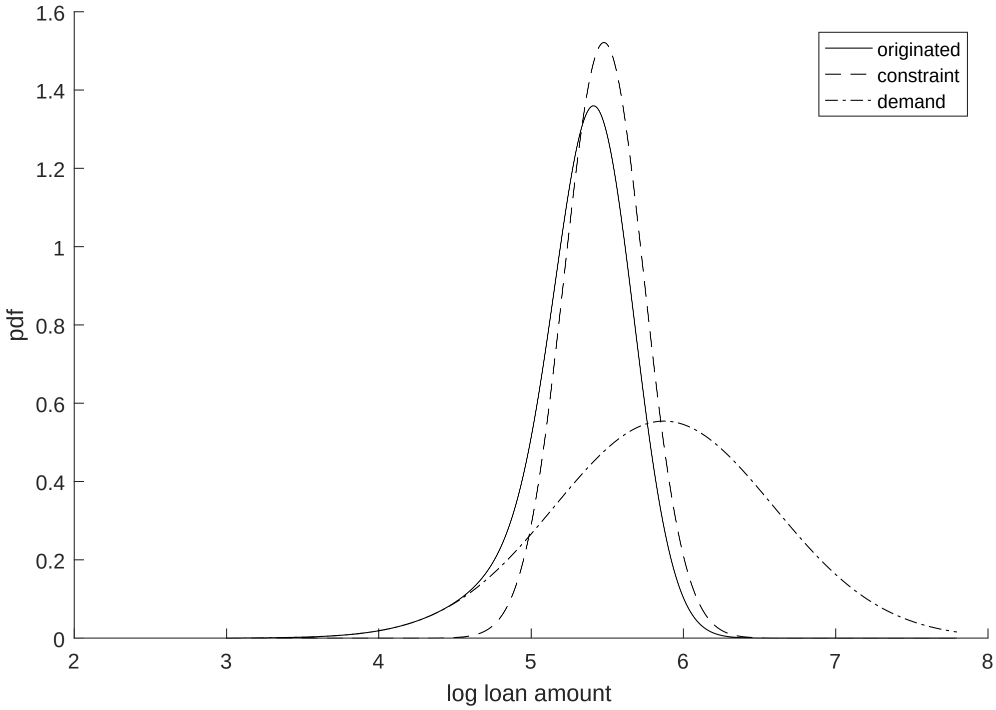


(b) Mortgage Credit Availability Index



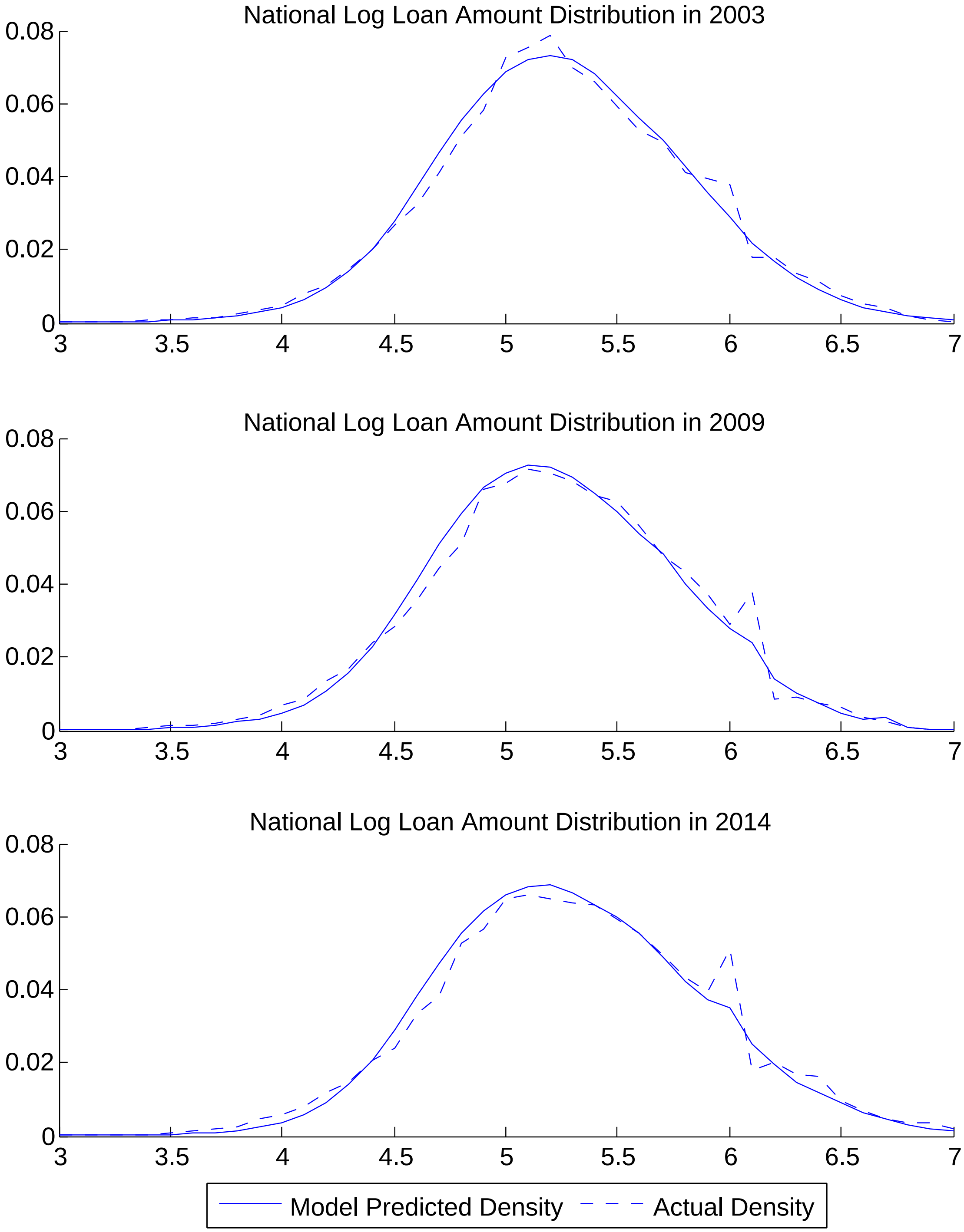
Panel (a) shows the correlation between the net fraction of banks reporting a tightening of standards for residential mortgages in the Senior Loan Officer Opinion Survey (SLOOS) and changes in the aggregate loan frontier. SLOOS responses are reported separately for prime, nontraditional and subprime loans. To obtain aggregate SLOOS responses for each year, we average three categories using equal weights. Also, we average quarterly responses to obtain annual estimates. Panel (b) shows the loan frontier along with the Mortgage Credit Availability Index (MCAI) produced by the Mortgage Bankers' Association. The MCAI is a function of the number of loan programs offered by large investors and the risk characteristics that define the types of loans that these programs will accept. The loan frontier is aggregated over metro areas, incomes, and downpayments using the weights described in Section 4.

Figure A10: Distribution of the Minimum of a Bivariate Normal



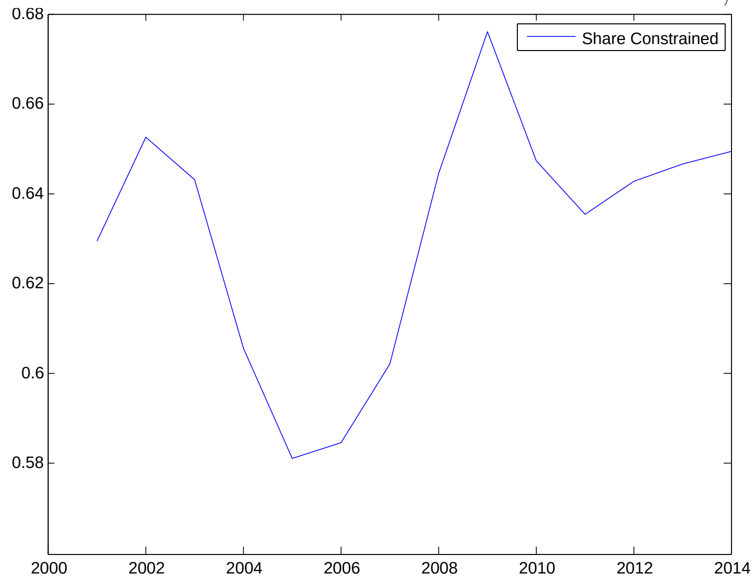
Constraints and unconstrained demand are jointly log-normal with parameters described in Section 5. Originated loan amount is the minimum of constraint and unconstrained demand.

Figure A11: Model Fit of Loan Amount Distributions, by Year



The figure compares the empirical distribution of mortgage originations with the simulated distribution from the estimates of the parametric model in Section 5.

Figure A12: Estimates for the Share of Constrained Borrowers, 2001-2014



The share of constrained borrowers is constructed by simulating the parametric model in Section 5 and computing the share of borrowers for whom the unconstrained borrowing demand is higher than their borrowing constraint.

References

Avery, Robert, Glen Canner, Ken Brevoort, and Neil Bhutta, "The 2011 Home Mortgage Disclosure Act Data," *Federal Reserve Bulletin*, 2012, *100* (4).

Barakova, Irina, Paul S. Calem, and Susan M. Wachter, "Borrowing constraints during the housing bubble," *Journal of Housing Economics*, 2014, *24*, 4–20.

Blackburn, McKinley L and Todd Vermilyea, "The prevalence and impact of misstated incomes on mortgage loan applications," *Journal of Housing Economics*, 2012, *21* (2), 151–168.

Fuster, Andreas and Basit Zafar, "The Sensitivity of Housing Demand to Financing Conditions: Evidence from a Survey," Working Paper, Federal Reserve Bank of New York 2015.