

Internet Appendix for
“Big Banks, Household Credit Access, and
Intergenerational Economic Mobility”

Erik J. Mayer*

Wisconsin School of Business
University of Wisconsin-Madison

*Contact: erik.mayer@wisc.edu, Grainger Hall, 975 University Avenue, Madison, WI 53706.

Internet Appendix A — Supplementary Figures and Tables

Figure IA1

Mortgage Approval and Borrower-Lender Distance at Small vs. Large Banks

This figure shows the relationship between mortgage approval rates and the distance from the property to the bank's nearest branch for applications received by commercial banks from 2010 to 2015. The left panel presents the results for low-income applicants (below the median U.S. household income), and the right panel shows the results for high-income applicants. The plots show approval rates at small banks (assets less than 1 Billion in 2010 dollars), and large banks (all other banks). The sample consists of all mortgage applications intended for home purchase in the Home Mortgage Disclosure Act database, excluding non-conventional applications (e.g., FHA, VA) and applications for loan amounts above the limits set for securitization by the Government Sponsored Enterprises (i.e., "jumbo loans"). I also require the property to be located within 20 miles of the bank's nearest branch and in a Metropolitan Statistical Area (where HMDA reporting is most comprehensive). The distance from the property to the bank's nearest branch is computed using the Haversine formula which gives the distance between two sets of longitude and latitude coordinates. The coordinates of the property are defined as the centroid of the census tract it is in, and the coordinates for bank branches are available from the Federal Deposit Insurance Corporation.

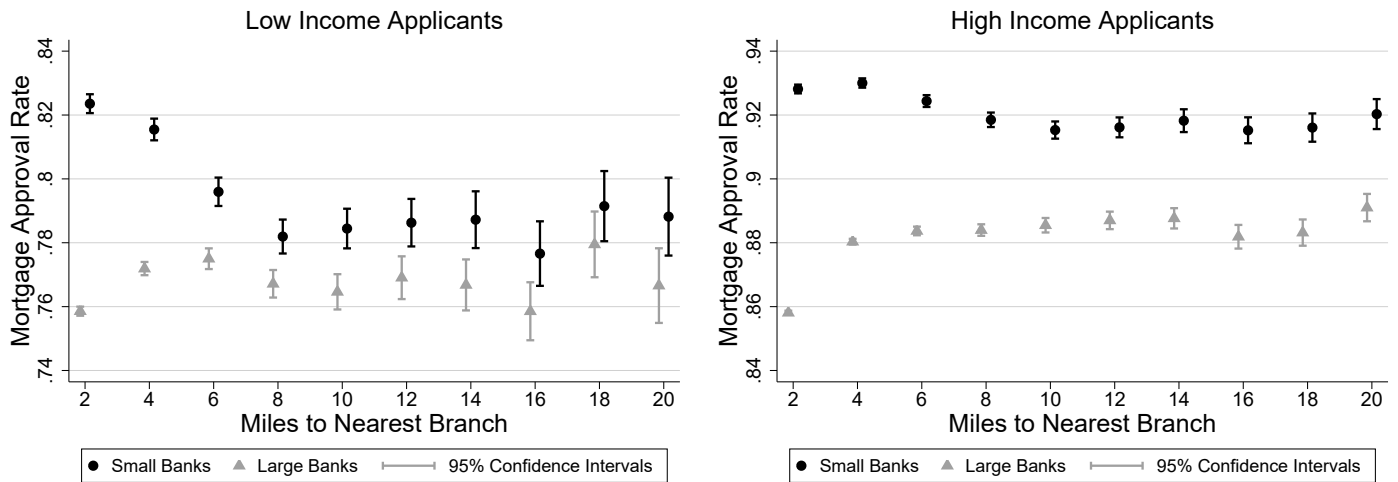


Figure IA2
Mortgage Approval at Small vs. Large Banks Over Time

This figure shows mortgage approval rates at small and large banks from 1995 to 2015. Small banks are those with assets less than 1 Billion in 2010 dollars, and the remaining banks are classified as large. The sample consists of all first-lien mortgage applications intended for home purchase in the Home Mortgage Disclosure Act database, excluding non-conventional applications (e.g., FHA, VA) and applications for loan amounts above the limits set for securitization by the Government Sponsored Enterprises (i.e., “jumbo loans”). Prior to 2004, second-lien “piggy-back” loans cannot be directly identified in HMDA, so I exclude applications in 2003 and earlier for loan amounts of less than \$40,000 (likely piggy-back loans). I also require the property to be in a Metropolitan Statistical Area, where HMDA reporting is most comprehensive. High/Low income applicants are defined as those with incomes above/below the U.S. median household income.

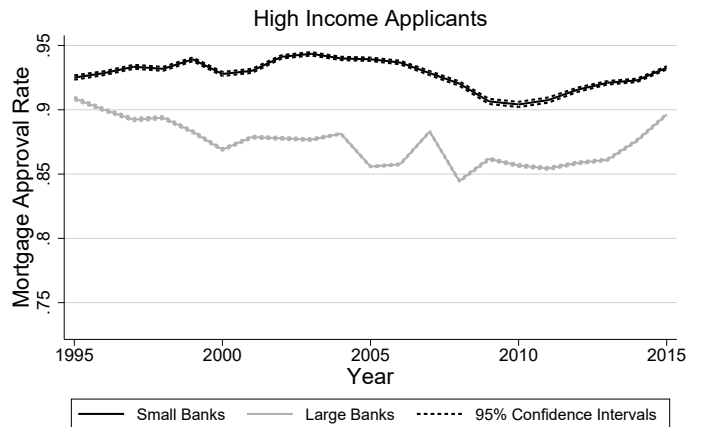
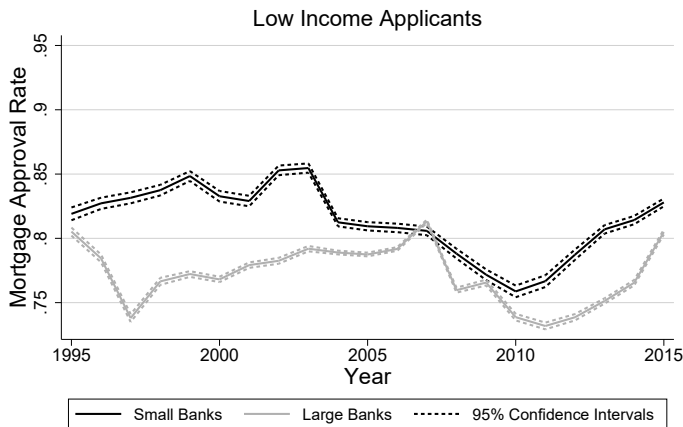
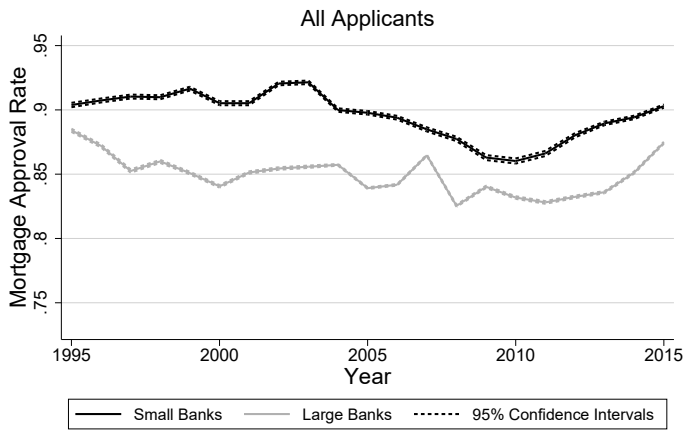


Table IA1: Interstate Banking Deregulation Years

This table presents the years that each state opened its borders to interstate banking by allowing interstate bank mergers.

State	Deregulation Year
Maine	1978
Alaska	1982
Connecticut	1983
Massachusetts	1983
Utah	1984
Kentucky	1984
Rhode Island	1984
North Carolina	1985
Nevada	1985
Virginia	1985
Idaho	1985
Ohio	1985
Georgia	1985
Tennessee	1985
Maryland	1985
District of Columbia	1985
Florida	1985
Minnesota	1986
New Jersey	1986
Michigan	1986
Missouri	1986
New York	1986
South Carolina	1986
Indiana	1986
Arizona	1986
Oregon	1986
Pennsylvania	1986
Illinois	1986
Wisconsin	1987
Texas	1987
Oklahoma	1987
Wyoming	1987
Louisiana	1987
Alabama	1987
New Hampshire	1987
California	1987
Washington	1987
South Dakota	1988
Colorado	1988
West Virginia	1988
Vermont	1988
Delaware	1988
Mississippi	1988
New Mexico	1989
Arkansas	1989
Nebraska	1990
Iowa	1991
North Dakota	1991
Kansas	1992
Montana	1993
Hawaii	1997

Table IA2: State Borders Where States Have a Large Contrast in Interstate Branching Policies

This table presents the state borders where the two states have a strong contrast in policies towards interstate bank branching as of the start of 2010. Columns 1 and 2 present the state with strong restrictions towards interstate branching and the value of the branching restrictions index developed in Rice and Strahan (2010). Columns 3 and 4 present the bordering state with fewer restrictions on interstate bank branching and its value of the restrictions index.

State with Strong Branching Restrictions	Restrictions Index	State Open to Entry	Restrictions Index
Alabama	3	Tennessee	1
Arkansas	4	Oklahoma	1
Arkansas	4	Tennessee	1
Colorado	4	Oklahoma	1
Colorado	4	Utah	1
Delaware	3	Maryland	0
Delaware	3	New Jersey	1
Delaware	3	Pennsylvania	0
Georgia	3	North Carolina	0
Georgia	3	Tennessee	1
Idaho	3	Utah	1
Idaho	3	Washington	1
Iowa	4	Illinois	0
Kansas	4	Oklahoma	1
Kentucky	3	Illinois	0
Kentucky	3	Indiana	1
Kentucky	3	Ohio	0
Kentucky	3	Tennessee	1
Kentucky	3	Virginia	0
Kentucky	3	West Virginia	1
Minnesota	3	Michigan	0
Minnesota	3	North Dakota	1
Mississippi	4	Tennessee	1
Missouri	4	Illinois	0
Missouri	4	Oklahoma	1
Missouri	4	Tennessee	1
Montana	4	North Dakota	1
Nevada	3	Utah	1
New Mexico	3	Oklahoma	1
New Mexico	3	Utah	1
Oregon	3	Washington	1
South Carolina	3	North Carolina	0
South Dakota	3	North Dakota	1
Wisconsin	3	Illinois	0
Wisconsin	3	Michigan	0
Wyoming	3	Utah	1

Table IA3: Large Bank Market Share and Household Credit Access: Heckman Estimation

This table repeats the baseline tests from Table 2, except using Heckman (1979) estimation methods. *Lambda* represents the inverse Mills ratio of the first-stage probit predicting sample inclusion. Similar to Table 2, the outcome variable in these second-stage tests is *Credit Approval*, which is defined for person-years where the individual applies for credit, and indicates whether they successfully opened a new credit line. The main explanatory variable is *Large Bank Market Share*, which is the fraction of bank branches located within 10 miles of the household that are owned by banks with more than \$1 Billion in assets (2010 dollars). The tests also control for individual, census tract, and county-level characteristics. Column 1 presents the effect of *Large Bank Market Share* on *Credit Approval* for all applicants. Columns 2, 3, and 4 interact *Large Bank Market Share* with indicators for the applicant having a low income, subprime credit score, or limited credit history, respectively. The base terms for these indicators are also included, and the tests use credit bureau data from 2010 to 2015. All continuous explanatory variables are standardized to have a mean of zero and a standard deviation of one, and coefficients are reported in percentage point units.

	(1)	(2)	(3)	(4)
Large Bank Market Share	-0.820*** (0.0222)	-0.446*** (0.0280)	-0.332*** (0.0280)	-0.459*** (0.0261)
Large Bank Market Share X Low Income		-0.785*** (0.0356)		
Large Bank Market Share X Subprime			-1.084*** (0.0364)	
Large Bank Market Share X Limited Credit History				-1.231*** (0.0375)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes
<i>Lambda</i>	-15.33*** (0.611)	-15.54*** (0.606)	-15.34*** (0.608)	-3.609*** (0.618)
<i>Prob > χ^2</i>	0.000	0.000	0.000	0.000
Observations	13,568,367	13,568,367	13,568,367	13,568,367
Censored	7,365,486	7,365,486	7,365,486	7,365,486

Table IA4: Robustness Tests Using ZIP Code Banking Markets

This table presents robustness tests that replicate the main results from Table 2 using two alternate approaches to define banking market shares. The first alternate approach is to construct *Large Bank Market Share* based on ZIP code deposit market shares using data from the FDIC Summary of Deposits. The second approach is to construct *Large Bank Market Share* using ZIP code bank branch counts. Panel A replicates the main results from Table 2 showing that *Large Bank Market Share* reduces credit approval, especially for subprime borrowers, where *Large Bank Market Share* is defined based on deposits in the ZIP code (columns 1 and 2) and branches in the ZIP code (columns 3 and 4). Coefficients are reported in percentage point units, and the standard errors are clustered by ZIP code-year. Panel B presents the median number of bank branches, large bank market share (based on branches), and bank branch HHI, when the local banking market is classified based on a 10-mile radius around households versus their ZIP code or county, respectively.

Panel A: Replication of Baseline OLS Results				
	LBMS Using ZIP Deposits		LBMS Using ZIP Branches	
	(1)	(2)	(3)	(4)
Large Bank Market Share	-0.189*** (0.0219)	0.0392 (0.0251)	-0.168*** (0.0223)	0.0594** (0.0254)
Large Bank Market Share X Subprime		-0.503*** (0.0368)		-0.506*** (0.0368)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes
R^2	0.193	0.193	0.193	0.193
Observations	5,769,101	5,769,101	5,769,101	5,769,101
Panel B: Additional Statistics on Banking Markets				
	10 Miles	ZIP Code	County	
Number of Branches	85	7	130	
Large Bank Market Share	0.850	0.869	0.847	
Bank Branch HHI	0.100	0.200	0.095	

Table IA5: Robustness Tests Using Alternate Large Bank Definitions

This table presents robustness tests that replicate the main results from Tables 2 and 6 using two alternate approaches to define the large/small bank variables. The first alternate approach is to use a \$10 Billion cutoff to delineate large versus small banks (rather than \$1 Billion). The second alternate approach is to exclude the four largest U.S. banks (JP Morgan Chase, Bank of America, Wells Fargo, and Citibank). Panel A presents tests that replicate the main results from Table 2 showing that *Large Bank Market Share* reduces credit approval, especially for subprime borrowers. Columns 1 and 2 use the \$10 Billion cutoff, and columns 3 and 4 exclude the top four banks. Panel B presents tests that replicate the main results from Table 6 showing that borrower-lender distance reduces mortgage approval, especially at small banks, and that small banks approve more applications overall. Again, columns 1 and 2 use the \$10 Billion cutoff, and columns 3 and 4 exclude the top four banks. Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year in Panel A and county-year in Panel B.

Panel A: Replication of Baseline OLS Results				
	\$10B Cutoff		Excluding Top 4 Banks	
	(1)	(2)	(3)	(4)
Large Bank Market Share	-0.545*** (0.0295)	-0.230*** (0.0318)	-0.514*** (0.0277)	-0.226*** (0.0294)
Large Bank Market Share X Subprime		-0.685*** (0.0359)		-0.649*** (0.0354)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes
R^2	0.194	0.194	0.194	0.194
Observations	6,202,881	6,202,881	6,202,881	6,202,881
Panel B: Replication of HMDA Mortgage Approval Results				
	\$10B Cutoff		Excluding Top 4 Banks	
	(1)	(2)	(3)	(4)
Distance To Branch	-0.0975*** (0.0102)	-0.0892*** (0.0166)	-0.0884*** (0.0110)	-0.101*** (0.0279)
Distance To Branch X Small Bank		-0.0543*** (0.0174)		-0.0617** (0.0279)
Small Bank		1.499*** (0.166)		1.959*** (0.180)
Applicant and Loan Characteristics	Yes	Yes	Yes	Yes
Census Tract Ratios and Averages	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes
County X Year FE	Yes	Yes	Yes	Yes
R^2	0.082	0.082	0.091	0.092
Observations	3,619,343	3,619,343	2,299,720	2,299,720

Table IA6: Large Bank Market Share and Household Credit Access: OLS Results for Subsamples

This table presents regressions of individuals' *Credit Approval* on *Large Bank Market Share* and individual, census tract, and county-level characteristics as well as state-by-year fixed effects. The full sample includes all person-years in the credit bureau dataset from 2010 to 2015 where the person applies for credit. *Credit Approval* is an indicator for the individual successfully opening a new credit line. *Large Bank Market Share* is the fraction of bank branches located within 10 miles of where the individual lives that are owned by banks with greater than \$1 Billion in assets (2010 dollars). The tests in this table split the sample based on characteristics of the location where people live. Columns 1 and 2 split the sample into urban and rural areas (in a metropolitan statistical area or not). Columns 3 and 4 split the sample based on income inequality (above/below median gini coefficient from county-level data published by Chetty et al. (2014)). Columns 5 and 6 split the sample based on the minority population share in the individual's census tract (above/below median). In each test, I interact *Large Bank Market Share* with *Subprime*, an indicator for the applicant having a *Credit Score* ≤ 660 . The base term for the interaction (*Subprime*) is omitted because it is a direct linear combination of the indicators for each 10 point bin of *Credit Score*. All continuous explanatory variables are standardized to have a mean of zero and a standard deviation of one. Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year.

	Urban / Rural Split		Income Inequality Split		Minority Share Split	
	MSA (1)	Non-MSA (2)	High (3)	Low (4)	High (5)	Low (6)
Large Bank Market Share	-0.172*** (0.0380)	0.364*** (0.0481)	-0.419*** (0.0583)	-0.00248 (0.0327)	-0.397*** (0.0556)	0.155*** (0.0331)
Large Bank Market Share X Subprime	-0.358*** (0.0504)	-0.852*** (0.0691)	-0.873*** (0.0689)	-0.721*** (0.0437)	-0.914*** (0.0651)	-0.913*** (0.0459)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.193	0.200	0.191	0.191	0.195	0.174
Observations	5,413,716	789,165	3,079,713	3,122,870	3,124,659	3,078,222

Table IA7: IV/2SLS Analysis of the Effect of Large Bank Market Share on Additional Measures of Credit Access

This table presents OLS and IV regressions of outcome variables describing a person’s access to credit on *Large Bank Market Share* and individual, census tract, and county-level characteristics as well as state-by-year fixed effects. *Large Bank Market Share* is the fraction of bank branches within 10 miles of the household that are owned by banks with greater than \$1 Billion in assets. I select all person-years in the credit bureau data from 2010 to 2015 where the person lives within 50 miles of a state border where there is a large contrast in the two states’ interstate bank branching policies (see Table IA2 for a list of these state borders). I instrument for *Large Bank Market Share* using a person’s *Position Relative to Border*, which ranges from -50 to 50, with -50 and 50 representing census tracts 50 miles towards the interior of the state with strong or weak branching restrictions, respectively. I also interact *Large Bank Market Share* with indicators for the borrower having a low income, subprime credit score, or limited credit history. The IV regressions instrument for these interactions with the interaction between the indicator and *Position Relative to Border*. The base terms for these interactions are omitted because they are direct linear combinations of the bin indicators used to control for the direct effect (i.e., bin indicators for *Estimated Income*, *Credit Score*, and *Number of Credit Lines*).

Panel A presents the results when *Credit Approval (First Mortgage)* is the outcome of interest. This variable measures mortgage credit approval for individuals with no outstanding mortgage at the start of the year, who attempt to open a mortgage during the year. Panels B and C present the results when an individual’s share of total debt held on credit cards, and an indicator for having retail debt, are the dependent variables. All continuous explanatory variables are standardized to have a mean of zero and a standard deviation of one. Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year.

Panel A: Credit Approval (First Mortgage)								
	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Large Bank Market Share	0.310*	0.892***	0.904***	0.441**	-1.581	-0.344	-0.603	-1.165
	(0.171)	(0.226)	(0.231)	(0.214)	(1.365)	(1.510)	(1.528)	(1.473)
Large Bank Market Share X Low Income		-1.069***				-2.243**		
		(0.233)				(1.088)		
Large Bank Market Share X Subprime			-1.183***				-1.989*	
			(0.228)				(1.073)	
Large Bank Market Share X Limited Credit History				-0.284				-0.924
				(0.222)				(1.032)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.168	0.168	0.168	0.168	-	-	-	-
Observations	129,555	129,555	129,555	129,555	129,555	129,555	129,555	129,555
First-stage F-stat	-	-	-	-	541.0	270.5	269.1	268.3
Panel B: Credit Card Debt Share								
	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Large Bank Market Share	0.319***	-0.366***	-0.0771*	-0.105**	1.403***	-0.957***	0.245	-0.424
	(0.0393)	(0.0430)	(0.0440)	(0.0407)	(0.316)	(0.320)	(0.323)	(0.317)
Large Bank Market Share X Low Income		1.629***				5.669***		
		(0.0523)				(0.284)		
Large Bank Market Share X Subprime			1.232***				3.545***	
			(0.0512)				(0.269)	
Large Bank Market Share X Limited Credit History				1.394***				6.133***
				(0.0609)				(0.333)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.320	0.321	0.321	0.321	-	-	-	-
Observations	1,809,982	1,809,982	1,809,982	1,809,982	1,809,982	1,809,982	1,809,982	1,809,982
First-stage F-stat	-	-	-	-	887.7	444.8	446.6	444.3

Panel C: Have Retail Debt

	OLS				IV			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Large Bank Market Share	-0.0287 (0.0376)	-0.129** (0.0510)	-0.439*** (0.0457)	-0.393*** (0.0514)	0.259 (0.313)	-0.0791 (0.355)	-1.598*** (0.333)	-0.204 (0.358)
Large Bank Market Share X Low Income		0.177*** (0.0496)				0.586** (0.251)		
Large Bank Market Share X Subprime			0.963*** (0.0471)				2.613*** (0.242)	
Large Bank Market Share X Limited Credit History				0.731*** (0.0485)				1.098*** (0.238)
Individual, Census Tract, and County Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State X Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R ²	0.330	0.330	0.330	0.330	-	-	-	-
Observations	2,541,947	2,541,947	2,541,947	2,541,947	2,541,947	2,541,947	2,541,947	2,541,947
First-stage F-stat	-	-	-	-	866.0	433.2	434.5	433.1

Table IA8: Estimating Differences Within Lender-by-Neighborhood Pairs

This table presents regressions of an indicator for a mortgage application being approved on the interaction term between indicators for the borrower having a low income (below the median U.S. household income) and for the bank being small (assets less than 1 Billion in 2010 dollars), as well as control variables. Note that the base terms for this interaction are subsumed by the income centile indicators and the fixed effects. The sample includes all first-lien home purchase mortgage applications in the Home Mortgage Disclosure Act database that were received by commercial banks from 2010-2015. I exclude non-conventional applications (e.g., FHA and VA), and applications for amounts above the limits set for securitization by the Government Sponsored Enterprises (i.e., “jumbo loans”). I also require the property to be located within 20 miles of the bank’s nearest branch and in a Metropolitan Statistical Area (where HMDA reporting is most comprehensive). The continuous explanatory variables are standardized to have a mean of zero and a standard deviation of one. Coefficients are reported in percentage point units, and the standard errors are clustered by county-year.

	(1)	(2)
Low Income X Small Bank	0.336**	0.362**
	(0.152)	(0.170)
<i>Applicant and Loan Characteristics</i>		
Income Centile Indicators	Yes	Yes
Loan Amount Centile Indicators	Yes	Yes
Loan To Income Ratio	-3.825***	-3.166***
	(0.133)	(0.197)
Joint Application	-0.130***	-0.364***
	(0.0494)	(0.0586)
African American	-7.741***	-6.676***
	(0.167)	(0.200)
Hispanic	-4.278***	-3.630***
	(0.120)	(0.124)
<i>Census Tract Ratios and Averages</i>		
Income / Tract Income	0.231***	-0.548***
	(0.0632)	(0.107)
Loan To Income / Tract Loan To Income	-1.038***	-2.064***
	(0.0847)	(0.185)
Loan Amount / Tract Loan Amount	-0.946***	-2.875***
	(0.0579)	(0.120)
Average Vantage Score _{t-1}	1.153***	
	(0.0449)	
Lender X County X Year FE	Yes	No
Lender X Tract X Year FE	No	Yes
R^2	0.142	0.370
Observations	3,595,808	2,781,202

Table IA9: Distance, Bank Size, and Mortgage Approval: Alternate Time Periods

This table presents robustness tests that replicate the main results from Table 6 (which uses data from 2010 to 2015) during earlier time periods. Panel A presents the results for the pre-crisis period of 1995 to 2006. Panel B presents the results for the financial crisis period of 2007 to 2009. As in Table 6, the dependent variable in the regressions is an indicator for a mortgage application being approved. The independent variables and sample construction also closely follow the tests in Table 6, with three exceptions. First, the control for the average credit score in the census tract is not available in these earlier periods. Second, the control for applicant ethnicity (*Hispanic*) is not available prior to 2004. Third, secondary liens are not identified in HMDA prior to 2004, so I drop all applications for loans less than \$40,000 in 2003 or earlier, and applications to Conseco Bank Inc., a lender specializing in piggy-back second-liens in the late 1990s. Coefficients are reported in percentage point units, and the standard errors are clustered by county-year.

Panel A: Pre-Crisis (1995-2006)					
	Full Sample			Low Income	High Income
	(1)	(2)	(3)	(4)	(5)
Distance To Branch	-0.174*** (0.0104)	-0.143*** (0.0147)	-0.108*** (0.0109)	-0.289*** (0.0189)	-0.0824*** (0.0151)
Distance To Branch X Small Bank		-0.0721*** (0.0171)		-0.104*** (0.0259)	-0.0661*** (0.0172)
Small Bank		1.313*** (0.127)		1.217*** (0.210)	1.367*** (0.119)
Distance To Branch X Low Income			-0.261*** (0.0123)		
Applicant and Loan Characteristics	Yes	Yes	Yes	Yes	Yes
Census Tract Ratios	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
County X Year FE	Yes	Yes	Yes	Yes	Yes
R ²	0.073	0.074	0.074	0.094	0.047
Observations	6,205,164	6,205,164	6,205,164	1,647,402	4,557,662
Panel B: Financial Crisis (2007-2009)					
	Full Sample			Low Income	High Income
	(1)	(2)	(3)	(4)	(5)
Distance To Branch	-0.155*** (0.0161)	-0.124*** (0.0205)	-0.0931*** (0.0175)	-0.235*** (0.0281)	-0.0934*** (0.0223)
Distance To Branch X Small Bank		-0.0568* (0.0299)		-0.0332 (0.0455)	-0.0576* (0.0308)
Small Bank		0.750*** (0.237)		0.654* (0.356)	0.998*** (0.232)
Distance To Branch X Low Income			-0.195*** (0.0244)		
Applicant and Loan Characteristics	Yes	Yes	Yes	Yes	Yes
Census Tract Ratios	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes
County X Year FE	Yes	Yes	Yes	Yes	Yes
R ²	0.079	0.079	0.079	0.106	0.057
Observations	2,048,123	2,048,123	2,048,123	565,880	1,482,231

Table IA10: Effect of Bank Size and Distance on Loan Sales and Rates

This table presents regressions where the dependent variable is an indicator for a mortgage being sold (columns 1-3) or an indicator for the mortgage having an APR more than 1.5 percentage points above the average prime rate and thus having the rate spread reported in HMDA (columns 4-6). The key independent variables are indicators for the bank being small, or for the borrower having a low income, as well as the distance from the property to the bank's nearest branch, and interactions. The sample includes all first-lien home purchase mortgages in the HMDA database that were originated by commercial banks from 2010 to 2015. I exclude non-conventional mortgages (e.g., FHA and VA), and loans for amounts above the limits set for securitization by the Government Sponsored Enterprises (i.e., "jumbo loans"). I also require the property to be located within 20 miles of the bank's nearest branch and in a Metropolitan Statistical Area (where HMDA reporting is most comprehensive). All specifications include the full set of control variables from the credit approval tests in Table 6, as well as county-by-year fixed effects. Coefficients are reported in percentage point units, and the standard errors are clustered by county-year.

	Dep. Var. = I(Loan Sold)			Dep. Var. = I(High APR)		
	(1)	(2)	(3)	(4)	(5)	(6)
Small Bank	-11.67*** (0.514)	-12.32*** (0.587)	-11.89*** (0.509)	0.912*** (0.189)	0.0521 (0.228)	0.771*** (0.190)
Distance To Branch		0.138*** (0.0312)	0.129*** (0.0291)		0.0503*** (0.0110)	0.0484*** (0.0104)
Distance To Branch X Small Bank		0.127** (0.0547)			0.143*** (0.0238)	
Distance To Branch X Low Income			0.171*** (0.0300)			0.232*** (0.0178)
Borrower and Loan Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Census Tract Ratios and Averages	Yes	Yes	Yes	Yes	Yes	Yes
Bank Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
County X Year FE	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.154	0.154	0.154	0.151	0.151	0.151
Observations	2,928,249	2,928,249	2,928,249	2,928,249	2,928,249	2,928,249

Table IA11: First Stage Regression of Large Bank Market Share on Years Since Interstate Deregulation

This table presents the first stage regression for the IV/2SLS analysis that estimates the effect of the share of large bank branches in a county on intergenerational mobility. The dependent variable in these first stage regressions is the share of large bank branches in a county. The instrumental variable is the years since the state removed its regulations preventing interstate bank mergers (*Years Since Interstate Deregulation*). The regressions also control for a broad set of county-level characteristics measured as of the year 2000. The sample includes the cross section of U.S. counties for which data on all of the covariates are available. Column 1 presents the results for the full sample. The remaining columns show the results for each subsample used in the second stage. Specifically, columns 2-5 show sample splits for counties above/below the median levels of affordable owner-occupied homes per capita, and student-teacher ratios at local schools. All explanatory variables are standardized to have a mean of zero and a standard deviation of one (except for *Years Since Interstate Deregulation*, which is in years), and the standard errors are robust.

	Full Sample	Affordable Homes		Student-Teacher Ratio	
	(1)	High (2)	Low (3)	High (4)	Low (5)
Years Since Interstate Deregulation	0.0843*** (0.00843)	0.0609*** (0.0108)	0.106*** (0.0137)	0.0815*** (0.0137)	0.0829*** (0.0116)
<u>Race and Segregation</u>					
Black Population Share	-0.0865** (0.0386)	-0.0501 (0.0508)	-0.107* (0.0552)	-0.0352 (0.0568)	-0.144*** (0.0548)
Racial Segregation	0.122*** (0.0218)	0.0826** (0.0346)	0.115*** (0.0291)	0.125*** (0.0293)	0.126*** (0.0337)
Segregation of Poverty	0.156*** (0.0250)	0.150*** (0.0446)	0.167*** (0.0297)	0.192*** (0.0310)	0.0823** (0.0400)
Commute Less Than 15min	0.152*** (0.0289)	0.123*** (0.0367)	0.199*** (0.0450)	0.217*** (0.0449)	0.104*** (0.0371)
<u>Income and Inequality</u>					
Per Capita Income	0.215*** (0.0279)	0.293*** (0.0542)	0.219*** (0.0379)	0.229*** (0.0434)	0.232*** (0.0368)
Gini Coefficient	0.00293 (0.0361)	0.0554 (0.0547)	-0.0253 (0.0494)	0.00190 (0.0519)	-0.0142 (0.0508)
Top 1 Percent Income Share	-0.00654 (0.0293)	-0.0337 (0.0442)	0.0101 (0.0408)	-0.00457 (0.0416)	0.00203 (0.0409)
<u>Family Characteristics</u>					
Single Mother Households	0.00671 (0.0503)	-0.107 (0.0701)	0.0701 (0.0733)	-0.0550 (0.0731)	0.122* (0.0709)
Fraction of Adults Divorced	-0.0369 (0.0237)	-0.0231 (0.0311)	-0.0411 (0.0366)	0.0162 (0.0363)	-0.102*** (0.0317)
Fraction of Adults Married	-0.0603* (0.0317)	-0.191*** (0.0565)	0.00783 (0.0389)	-0.0346 (0.0448)	-0.0788* (0.0448)
<u>K-12 Education</u>					
K12 Student Teacher Ratio	0.190*** (0.0210)	0.169*** (0.0353)	0.148*** (0.0288)	0.191*** (0.0377)	0.145*** (0.0529)
K12 Test Scores (Income Adjusted)	-0.00482 (0.0226)	0.0303 (0.0330)	-0.0268 (0.0323)	-0.0300 (0.0329)	0.0487 (0.0339)
<u>Social Capital</u>					
Social Capital Index	-0.0686*** (0.0262)	-0.0454 (0.0376)	-0.110*** (0.0394)	-0.135*** (0.0402)	0.000970 (0.0354)
Religious Population Share	-0.139*** (0.0224)	-0.185*** (0.0271)	-0.0544 (0.0377)	-0.0512 (0.0342)	-0.222*** (0.0277)
Violent Crimes Per Capita	0.0388* (0.0201)	0.0225 (0.0297)	0.0621** (0.0283)	0.0263 (0.0267)	0.0553* (0.0320)
<u>Additional Controls</u>					
HHI of Bank Branches	-0.104 (0.158)	0.245 (0.206)	-0.314 (0.221)	0.169 (0.260)	-0.244 (0.189)
Log(Population Density)	0.0838** (0.0394)	0.210*** (0.0572)	0.0292 (0.0549)	0.0593 (0.0566)	0.171*** (0.0555)
Per Capita Income Growth (1980-2005)	-0.0642*** (0.0214)	-0.0106 (0.0262)	-0.103*** (0.0351)	-0.125*** (0.0359)	-0.0236 (0.0266)
R^2	0.369	0.345	0.328	0.310	0.343
Observations	2,417	1,212	1,205	1,264	1,153

Table IA12: Large Banks and Homeownership Growth in the 1990s

This table presents IV/2SLS estimates of the effect of *Large Bank Market Share* in a county on the growth in homeownership during the 1990s, defined as the change in the homeownership rate between 1990 and 2000. The number of years since the state removed its barriers to interstate bank mergers (i.e., interstate deregulation) is used to instrument for *Large Bank Market Share*. The controls follow prior tests and the full sample consists of the cross section of U.S. counties for which data on all the covariates are available. Column 1 presents the results for the full sample. Columns 2 and 3 show sample splits for counties above/below the median level of affordable owner-occupied homes per capita. Explanatory variables are standardized to have a mean of zero and a standard deviation of one, and the standard errors are robust.

	Full Sample	Affordable Homes	
	(1)	High (2)	Low (3)
Large Bank Market Share	-0.00388 (0.00283)	-0.0105** (0.00471)	0.00134 (0.00391)
All Controls	Yes	Yes	Yes
Observations	2,558	1,314	1,244
First-stage F-stat	117.7	43.6	64.7

Internet Appendix B — Matching Credit Bureau and HMDA Data

This Appendix describes the process I use to match mortgages in the credit bureau data to the Home Mortgage Disclosure Act database. This match allows information on the originating lender in the HMDA data to be combined with information on loan performance from the borrower's credit bureau records. Both data sources are anonymized, and there is no unique identifier to link the two datasets. However, the information on originated mortgages is reported at such a granular level in both datasets that the majority of mortgages can be uniquely identified and matched based on their characteristics.

I focus on conventional home purchase mortgages for amounts below the “jumbo” threshold, and match loans in the credit bureau data to originated loans in the HMDA data based on the following characteristics: origination year, census tract location, loan amount, whether the mortgage is joint or belongs to a single borrower, and if/to which quasi-government entity the loan is sold. I drop observations that are not uniquely identified. Because the HMDA data contain more than 95% of all originated mortgages, requiring the HMDA mortgage to be unique ensures that any matching mortgage in the credit bureau data identifies the same borrower with near-certainty. Fortunately, 80.5% of conventional home purchase mortgages in the HMDA data are uniquely identified based on the matching variables.

After identifying unique loans in the HMDA data, I make several additional requirements to ensure the quality of the match. I require mortgages to be on owner-occupied homes, so that the property location will match the borrower's location in the credit bureau data. I also require the mortgage to be a first lien, and the property must be located within an MSA (where the HMDA data are most comprehensive).

I apply a similar set of filters to the mortgages from the credit bureau data. I require the mortgage to be the borrower's only first lien mortgage at the time. This filter ensures that the borrower's location in the credit bureau data will match the property location in the HMDA data. I also require that the person live in an MSA directly following the loan origination. After combining the filters imposed on the HMDA and credit bureau data, the target population for the matched

sample is borrowers taking out a conventional home purchase mortgage for an amount below the “jumbo” threshold, on their primary residence, which is located within an MSA, from 2010 to 2013.

There are two potential sources of matching error. First, a data error in one of the matching variables could create a mismatch, but such errors should be rare because institutions systematically report these data to both the HMDA database and the credit bureaus. A second type of error could occur if a HMDA-reporting lender, and a non-reporting lender, originate identical mortgages that are otherwise unique. The reporting lender’s loan could be matched to the credit bureau record of either of the two borrowers. This type of mismatch should be rare because HMDA covers nearly the universe of mortgages. Moreover, which credit bureau record the HMDA loan matches is random, because it will depend on which record is in the 1% random sample of credit bureau data. Therefore, this type of mismatch should not create any bias in the estimates outside of pure noise.

Table IB1 presents summary statistics on the match. Panel A presents the match rate, which shows that I find a matching HMDA mortgage for 61.2% of the mortgages in the credit bureau data. Because HMDA includes 95% of mortgages and 80.5% are unique, the best I can do is roughly 76.5% (0.95×0.805). In other words, the algorithm found matches in roughly 80.0% ($\frac{0.612}{0.765}$) of the cases it could have. I view this as a reasonably good match. I require exact matching (rather than nearest matches or propensity scores) in order to keep the matched dataset as accurate as possible.

[Insert Table IB1 Here]

The summary statistics in Panel B of Table IB1 examine whether the matched sample is representative of the original population. Panel B shows that the sample of successfully matched mortgages is broadly representative of the starting population of credit bureau mortgages. Although there are statistically significant differences between matched and unmatched credit bureau mortgages, these differences are typically relatively small in economic terms.

Figure IB1 provides additional validation of the match by showing that the loan amount distribution for the matched mortgages is nearly identical to the distribution for the starting

sample of credit bureau mortgages.

[Insert Figure IB1 Here]

The loan performance tests which motivate the credit bureau to HMDA match focus on mortgages originated by banks. Table IB2 presents summary statistics for the subset of matched loans that were originated by banks. This sample, which is ultimately used in the analysis, contains 30,954 conventional home purchase mortgages originated by banks from 2010 to 2013, where the property is within an MSA, and the loan amount is below the “jumbo” threshold.

[Insert Table IB2 Here]

Figure IB1
Loan Amounts for Credit Bureau Mortgages and the Subset that Matched to HMDA

This figure shows the loan amount distribution for the sample of mortgages in the credit bureau data and for the subset of these mortgages that successfully matched to HMDA data. The starting sample of credit bureau data contains conventional home purchase mortgages originated from 2010 to 2013, where the property is located within a metropolitan statistical area, and the loan amount is below the Government Sponsored Entities' securitization limits (excludes "jumbo" loans). Note that for most areas, this limit is \$417,000, but that it is higher in certain high-cost areas. The matching is done based on the loan origination year, the census tract of the property, the loan amount, whether the mortgage is joint or belongs to a single borrower, and whether the loan is purchased by Fannie Mae or Freddie Mac on the secondary market. Only the mortgages in the HMDA data that are unique based on these matching variables are used as potential matches.

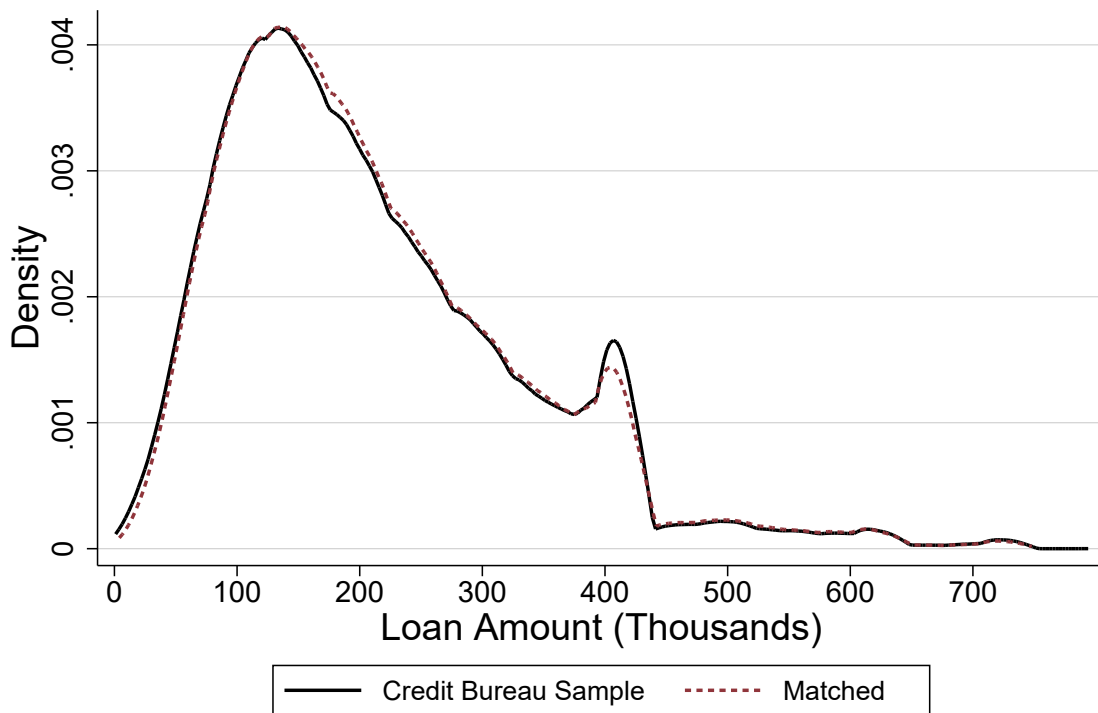


Table IB1: Summary of the Credit Bureau to HMDA Match

This table summarizes the match between mortgages in the credit bureau data and Home Mortgage Disclosure Act data. The starting sample of credit bureau data contains conventional home purchase mortgages originated from 2010 to 2013, where the property is located within a metropolitan statistical area, and the loan amount is below the Government Sponsored Entities' securitization limits (excludes "jumbo" loans). The matching is done based on the loan origination year, the census tract of the property, the loan amount, whether the mortgage is joint or belongs to a single borrower, and whether the loan is purchased by Fannie Mae or Freddie Mac on the secondary market. Only the mortgages in the HMDA data that are unique based on these matching variables are used as potential matches. Panel A shows the success rate of the matching approach. Panel B summarizes borrower, loan, and location characteristics for the mortgages in the credit bureau data, the subsample that were successfully matched to HMDA, and the unmatched loans. The final two columns show the normalized difference and the result of a t-test comparing the mean of the matched sample to the mean of the unmatched sample.

Panel A: Match Rate					
	Credit Bureau Sample	Matched to HMDA	Match Rate		
	114,330	69,930	61.17%		
Panel B: Summary Statistics for Matched vs. Unmatched					
	Credit Bureau Sample	Matched to HMDA	Unmatched	Matched vs. Unmatched	
	(N=114,330)	(N=69,930)	(N=44,400)	Norm. Diff.	t-stat
<i>Borrower and Loan Characteristics</i>					
Credit Score $t-1$	753.1	754.6	750.8	0.04	6.12
Number of Credit Lines $t-1$	6.605	6.453	6.845	-0.07	-9.53
Age	45.7	44.8	47.1	-0.12	-15.24
Total Debt $t-1$	120,536	108,487	139,722	-0.14	-10.21
Past Due Debt $t-1$	150.4	113.1	209.9	-0.02	-4.26
Joint Mortgage	0.531	0.526	0.539	-0.02	-3.67
Loan Amount	203,397	204,871	201,055	0.02	2.17
<i>County Characteristics</i>					
High School Diploma	0.873	0.877	0.868	0.12	3.09
Poverty	0.129	0.128	0.130	-0.04	-2.25
Minority Population Share	0.350	0.345	0.360	-0.05	-1.45
Unemployment Rate	0.081	0.079	0.083	-0.11	-3.68
Personal Income Per Capita	46,449	46,509	46,353	0.01	0.75

Table IB2: Summary Statistics for Mortgages Originated by Banks

This table presents summary statistics describing the subset of mortgages from the Credit Bureau/HMDA matched dataset that were originated by banks. The sample includes 30,954 conventional home purchase mortgages originated from 2010 to 2013. The property is required to be in a metropolitan statistical area, and the loan amount must be below the Government Sponsored Entities' securitization limits (excludes "jumbo" loans).

	Mean	Std. Dev.	P10	P50	P90
<i><u>Loan Outcomes</u></i>					
Delinquent (60 days)	0.035	0.184	0	0	0
<i><u>Borrower and Loan Characteristics</u></i>					
Credit Score t_{-1}	755.4	59.8	667.0	769.0	818.0
Number of Credit Lines t_{-1}	6.417	4.003	2.0	6.0	12.0
Age	44.772	13.918	28.0	43.0	64.0
Total Debt t_{-1}	108,949	147,063	757	41,273	295,501
Past Due Debt t_{-1}	104.7	2,019.8	0	0	0
Joint Mortgage	0.529	0.499	0	1	1
Loan Amount	204,507	121,308	76,000	176,000	380,000
Application Income	94,903	78,182	35,000	78,000	168,000
Loan To Income Ratio	2.479	1.340	1.148	2.339	4.000
Black	0.025	0.157	0	0	0
Hispanic	0.054	0.225	0	0	0
<i><u>Bank Characteristics</u></i>					
Equity Capital Ratio	0.110	0.021	0.085	0.112	0.131
Real Estate Loans Ratio	0.385	0.142	0.238	0.376	0.596
Profitability	0.007	0.011	-0.002	0.010	0.013