JOURNAL OF FINANCIAL AND QUANTITATIVE ANALYSIS

doi:10.1017/S0022109023001114

Big Banks, Household Credit Access, and Intergenerational Economic Mobility

Erik J. Mayer
University of Wisconsin–Madison Wisconsin School of Business
erik.mayer@wisc.edu

Abstract

Consolidation in the U.S. banking industry has led to larger banks. I find that low-income households face reduced access to credit when local banks are large. This result appears to stem from large banks' comparative disadvantage using soft information, which is particularly important for lending to low-income households. In contrast, the size of local banks has little or no effect on high-income households. Consistent with low-income parents' credit constraints limiting investment in their children's human capital, areas with larger banks exhibit a greater sensitivity of educational attainment to parental income, and less intergenerational economic mobility.

I. Introduction

The U.S. banking industry has experienced tremendous consolidation since states began removing barriers to bank expansion in the 1970s, leading to much larger banks. From the average U.S. household's perspective, the median-sized bank within 10 miles of their home is over 7 times larger today than it was in 1995. In this paper, I test whether the size of banks affects households' access to credit, and through this channel, intergenerational economic mobility.

It is unclear whether we should expect larger banks to lead to more or less credit access for households. Stein (2002) predicts small banks will have a comparative advantage using soft information to reduce information asymmetries, which should increase credit access (Stiglitz and Weiss (1981)). On the other hand, large banks benefit from economies of scale, and from diversification that reduces the cost of delegated monitoring (Diamond (1984)) and allows banks to lend out a higher proportion of their capital (Demsetz and Strahan (1997)). If these benefits outweigh any effects of reduced soft information utilization, we might expect larger

For helpful comments, I thank Alex Butler, Charles Calomiris, Mara Faccio (the editor), Gustavo Grullon, Christoffer Koch, Claire Labonne, Aaron Pancost, Amiyatosh Purnanandam, James Weston, and seminar participants at Rice University, the University of Michigan, Southern Methodist University, Vanderbilt University, the University of Houston, the University of Arizona, and the American University. I also thank conference participants at the 2020 Columbia/Banking Policy Institute Research Conference, the 2020 Community Banking in the 21st Century Research Conference, the 2019 Financial Management Association Meeting, the 2018 Lone Star Finance Conference, and the 2018 Banking and Finance Workshop at the Federal Reserve Bank of Dallas.

banks to improve households' access to credit, especially in cases where soft information is less important.

I find that borrowers of low economic status (i.e., low income, subprime credit score, and/or limited credit history) experience lower credit approval rates when local banks are large. In contrast, the size of banks has little or no effect on borrowers of high economic status. The evidence suggests that this asymmetric effect stems from the increased importance of soft information for lending to low-income households. These findings naturally raise the question of whether consolidation in the banking industry contributes to economic inequality.

I study intergenerational economic mobility, where theoretical models predict that credit access allows low-income households to invest in their children's human capital, leading to increased upward mobility across generations (e.g., Becker and Tomes (1979), (1986)). Specifically, I test whether having large local banks reduces intergenerational mobility due to the additional credit constraints low-income households face. I find evidence in support of this hypothesis using newly available data on mobility from Chetty, Hendren, Kline, and Saez (2014). This finding constitutes the first evidence of a link between the characteristics of financial institutions and intergenerational mobility.

The first set of empirical tests examine household credit access using a nationally representative sample of credit bureau records that provide individuals' age, census tract, credit score, debt by category (mortgage, auto, etc.), credit application inquiries, and other financial variables. The baseline OLS regressions show that LARGE_BANK_MARKET_SHARE (the fraction of bank branches within 10 miles of a borrower owned by banks with assets over $1 billion) has a negative effect on credit approval rates for borrowers of low economic status. These regressions control for borrower credit scores and individual, census tract, and county-level characteristics, as well as state-by-year fixed effects. Moreover, this result holds even in tests that isolate only the within-neighborhood time-series variation in bank characteristics that is due to banks entering and exiting local markets.

Despite the robust baseline results, this paper must address the challenges to identification that may arise from borrowers' credit applications not being randomly assigned to banks. The potential selection comes in two layers. First, borrowers can choose which banks to try when applying for credit. Importantly, we should expect this source of selection to work against finding that the composition of local banks affects credit access, because borrowers likely end up applying for credit at the type of bank most willing to lend to them (either knowingly, or through trial and error shopping).[1] The second layer of potential selection stems from the fact that the composition of local bank branches reflects banks' location decisions. If these decisions are correlated with an unobservable component of borrower credit quality, it could generate an omitted variables problem. However, it is important to point out that if large banks systematically build/buy branches where they want to lend to local households, we might expect

[1]I measure credit approval based on the success rate of annual credit shopping attempts rather than individual applications. This approach is conservative because it allows borrowers to shop for credit and potentially try multiple lenders.

the resulting bias to work *against* the baseline OLS result that large banks reduce households' access to credit.

To avoid an omitted variables bias, I employ an instrumental variables approach that isolates exogenous variation in LARGE_BANK_MARKET_SHARE. I exploit differences in state policies that restrict the ability of out-of-state banks (e.g., national banks) to enter local markets by building new branches or purchasing existing ones. I identify 36 state borders where one state has strong regulatory barriers to out-of-state bank branching, and the other state is open to entry. Unsurprisingly, branches in the state with barriers to out-of-state bank entry are owned by smaller banks. I select everyone in the credit bureau data living within 50 miles of these borders and use their location relative to the border to instrument for LARGE_ BANK_MARKET_SHARE. The differences in regulation make a person's position relative to the border an instrument for LARGE_BANK_MARKET_SHARE even when comparing two people living in the same state. For example, a person living 20 miles toward the interior of the state with regulatory barriers and small banks will have a lower LARGE_BANK_MARKET_SHARE than someone in the same state who lives near the border, because the neighboring state's banks are large.

The identifying assumption this approach makes is that for two borrowers in the same state during the same year, controlling for credit scores and individual, census tract, and county-level characteristics, their distance to the state border affects credit approval only through its effect on LARGE_BANK_MARKET_SHARE. The results from these instrumental variables tests show that a standard deviation increase in LARGE_BANK_MARKET_SHARE decreases subprime borrowers' overall credit approval rates by 3.7 percentage points compared to their mean approval rate of 53.0%, whereas the effect on prime borrowers' credit approval is positive and small in magnitude. The estimated effect of LARGE_BANK_MARKET_SHARE is larger in the instrumental variables regressions than their OLS counterparts, suggesting that any omitted variables bias indeed works against the OLS results.

The second set of empirical tests use Home Mortgage Disclosure Act (HMDA) data on mortgage applications, where the lender's identity is directly reported. I test the hypothesis that low-income households have better access to mortgage credit at small banks, and that small banks have a comparative advantage using soft information. I find that small banks approve a higher percentage of mortgage applications, consistent with these banks collecting soft information to price risks and ration credit less. I also find that as the distance from the property to the lender's nearest branch increases, approval rates decrease, especially when the borrower has a low income and/or the bank is small. Following the interpretation in the literature that borrower–lender distance affects credit provision through soft information production (e.g., Petersen and Rajan (2002), DeYoung, Glennon, and Nigro (2008), and Agarwal and Hauswald (2010)), these results indicate that soft information is especially important when lending to low-income households, and that smaller banks incorporate more of this information into lending decisions.

Consistent with an information advantage, small banks' mortgages have similar (or lower) delinquency rates compared to large banks', despite higher

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

approval rates. When I examine secondary market loan sales, I also find that small banks retain a much larger share of the mortgages they originate. Moreover, the likelihood that loans are sold increases with borrower–lender distance, especially at small banks, and for low-income borrowers.

After establishing that low-income households face tighter credit constraints when local banks are large, I test the hypothesis that large banks reduce intergenerational mobility. This third set of empirical tests use newly available mobility statistics computed at the county level by Chetty et al. (2014) from the IRS tax returns of children born in the early 1980s and their parents. Controlling for a broad set of covariates outlined in Chetty et al. (2014), plus additional controls, I find that the share of bank branches in a county owned by large banks has a negative effect on mobility levels.

I take two different, and complementary approaches to mitigate potential omitted variable concerns. First, I include state fixed effects in the regression, and the results remain virtually unchanged. This test shows that differences in various state policies outside of banking are not acting as omitted variables and driving the results. Second, I isolate plausibly exogenous variation in the size of local banks during the childhood of children in the Chetty et al. (2014) data based on states' staggered removal of regulations preventing interstate bank mergers from 1978 to 1997. Prior to deregulation, out-of-state banks could not enter local markets, making the number of years since the state deregulated a powerful instrument for large banks' market share.[2] Importantly, this instrumental variables approach exploits across-state variation, and therefore mitigates omitted variables that may operate within states. The instrumental variables results again show that having larger banks leads to lower intergenerational mobility levels (a standard deviation increase in the share of large bank branches in a county causes a reduction in intergenerational mobility of between 8% and 14%).[3]

Two further tests suggest that this effect is indeed driven by credit constraints limiting human capital formation in low-income households. First, tests show that the effect of banks on mobility is larger in areas where credit access is more likely to translate to homeownership (which has been shown to benefit children), and in areas with less government investment in children (where the financial burden falls more squarely on households). A second, fairly direct test, shows that having large local banks leads to children's human capital attainment (college attendance) being more sensitive to parental income, consistent with constraints limiting investment in children from low-income households.

This paper is related to studies showing small banks are important providers of credit to small businesses, consistent with an advantage lending based on soft information.[4] Although recent work suggests soft information matters when

[2]I follow a long literature in finance that treats these deregulation events as plausibly exogenous to local economic conditions (see Kroszner and Strahan (2014) for a review).

[3]As discussed in subsequent sections, I use two county-level measures of mobility from Chetty et al. (2014): the slope from a rank–rank regression of child income on parental income (percentile ranks from the national income distribution), and the probability that children from families in the bottom 40% of the income distribution transition out of this bottom 40% in adulthood.

[4]See, e.g., Strahan and Weston (1998), Cole, Goldberg, and White (2004), Berger, Miller, Petersen, Rajan, and Stein (2005), Carter and McNulty (2005), Berger and Black (2011), and Berger, Bouwman, and Kim (2017).

lending to households (e.g., Agarwal, Ambrose, Chomsisengphet, and Liu (2011), Iyer, Khwaja, Luttmer, and Shue (2016)), evidence on whether small banks play a special role in this setting is limited.[5] My paper contributes to this literature by providing loan-level evidence that small banks incorporate more soft information when lending to households, and by showing that low-income households are most affected by the size of local banks.

This paper is also connected to studies on the effects of banking deregulation and consolidation. These studies typically examine the effects on economic growth, or on firms.[6] Recent work also finds mixed evidence on the net effect of banking deregulation and consolidation on households' access to bank accounts (Bord (2018), Celerier and Matray (2019)). In contrast, my paper shows the effects on the distribution of credit across households, and ultimately on intergenerational economic mobility.

My work is also closely related to papers examining the effect of credit constraints on intergenerational mobility. Several studies using household survey data find that constraints reduce mobility (Gaviria (2002), Mazumder (2005)). However, Black and Devereux (2011) review this literature and point out that it relies on small samples and struggles to address endogeneity issues that arise from using wealth as a proxy for credit constraints. I contribute to this literature by providing evidence that low-income households' credit constraints reduce mobility using variation in banking deregulation and the size of local banks. This paper's findings also provide the first evidence of a link between the structure of the banking industry and intergenerational mobility.

II. Historical Restrictions on Bank Expansion

Banks in the United States have faced restrictions on geographic expansion since the Constitution gave states the right to charter and regulate banks (see Kroszner and Strahan (2014)). Prior to 1970, most states restricted even intrastate branching. Then, throughout the 1970s and 1980s, states removed these restrictions and allowed banks to build branches and convert subsidiaries and new acquisitions in their state into branches. This intrastate banking deregulation started the process of banking consolidation, by facilitating the formation of mid-sized community banks.

The states also historically used their authority to limit banks' expansion across state borders by prohibiting cross-state ownership of banks (interstate banking) and bank branches (interstate branching). The process of removing barriers to interstate banking began in 1978, when Maine decided to allow out-of-state banking companies to acquire its banks, as long as the acquirer's home state reciprocated

[5]Notably, Loutskina and Strahan (2011) show that banks operating primarily in one metropolitan area are more active in the jumbo mortgage market, consistent with an advantage using soft information.

[6]For the effect of deregulation on economic growth, see, e.g., Jayaratne and Strahan (1996) and Berger, Butler, Hu, and Zekhnini (2021). For the effects on firms, see, e.g., Cetorelli and Strahan (2006), Rice and Strahan (2010), and Chava, Oettl, Subramanian, and Subramanian (2013). For the effect of consolidation on small business lending, see Minton, Taboada, and Williamson (2021).

and gave banks in Maine the right to acquire banks in their state. Other states began to pass similar laws starting in 1982, and by 1993 every state except Hawaii allowed interstate banking (see Table IA1 in the Supplementary Material for the years that states deregulated). I use these staggered interstate deregulation events to isolate plausibly exogenous variation in the size of banks as of 1995 in order to study the effect of local banks' size on intergenerational mobility.

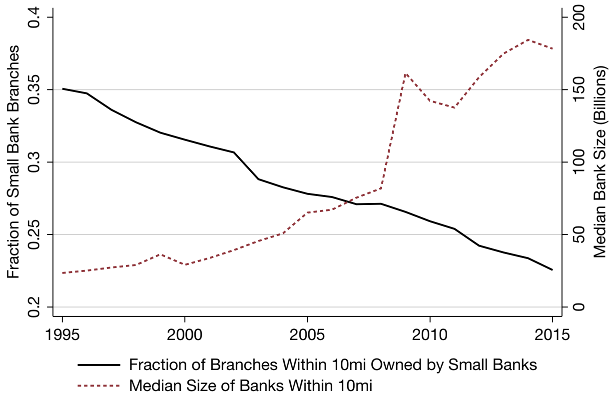
Although states opened their borders to bank acquisitions throughout the 1980s, only a few states allowed out-of-state banks to establish branches in their state prior to the passage of the Riegle-Neal Interstate Banking and Branching Efficiency Act (IBBEA) in 1994. The IBBEA removed remaining federal barriers and allowed bank holding companies (BHCs) to engage in interstate branching. However, the IBBEA also gave states the power to erect barriers to limit the entry of out-of-state banks through regulatory provisions. I follow Rice and Strahan (2010) and construct an index of states' policies to limit out-of-state bank entry. I then use 36 state borders where states have large differences in interstate branching policies to study the effect of the size of banks on household credit access using credit bureau data from 2010 to 2015 (see Table IA2 in the Supplementary Material for a list of these borders).

The steady removal of historical restrictions on bank expansion has led to significant consolidation in banking. Figure 1 describes banking consolidation from 1995 to 2015 from the average U.S. household's perspective. By examining all bank branches within 10 miles of households, the graph shows a sharp decline in the fraction of branches owned by small banks and a large increase in the median size of local banks.

FIGURE 1

Banking Consolidation from U.S. Households' Perspective

Figure 1 shows, from the average U.S. household's perspective, the fraction of local bank branches owned by small banks, and the median size of local banks. Local branches are defined as those within 10 miles of households, and small banks are those with less than 1 billion in assets in 2010 dollars. If a bank is owned by a holding company, the size of the bank is set as the combined size of all banks in the holding company. The location of households is set as the centroid of the census tract they live in, and the locations of bank branches are specific longitude and latitude coordinates from the Summary of Deposits available from the Federal Deposit Insurance Corporation. Distances between households and bank branches are computed based on longitude and latitude using the Haversine formula.



https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

III. Data and Methods

A. Data Sources Overview

To study households' access to credit, I use a nationally representative sample of credit bureau records, and Home Mortgage Disclosure Act data on mortgage applications. These data sets offer broad coverage to study approval rates on households' credit applications at a granular level, controlling for a rich set of covariates. Then, I evaluate the effect of local banks on intergenerational economic mobility, using county-level statistics on mobility and the sensitivity of educational attainment to parental income, published by Chetty et al. (2014). I discuss each of these three primary data sources in subsections below.

The main explanatory variables of interest in this paper are the characteristics of local banks, or of the specific bank receiving the credit application, when this is directly observable (i.e., when using HMDA data). The locations of bank branches in terms of latitude and longitude are available from the Summary of Deposits data published by the Federal Deposit Insurance Corporation (FDIC). I match bank branches to the commercial banks that own them, and collect data on these banks' characteristics from the Reports of Condition and Income (Call Reports) published by the Federal Financial Institutions Examination Council.

To control for a broad set of characteristics describing a location, I use census tract and county-level data from the U.S. Census Bureau. I also use county-level data on unemployment rates, personal income, and house prices from the Bureau of Labor Statistics, the Bureau of Economic Analysis, and the Federal Housing Finance Agency, respectively. The paper also uses additional county-level control variables collected from the National Center for Education Statistics, the George W. Bush Global Report Card, the Association of Religion Data Archives, and the Federal Bureau of Investigation. I also use county-level statistics describing income inequality computed in Chetty et al. (2014), and the county-level measure of social capital computed in Rupasingha, Goetz, and Freshwater (2006) as controls.

B. Credit Bureau Data

This paper uses a panel data set of anonymized individual credit bureau records. The data are a 1% representative sample of all U.S. residents with a credit history and Social Security number. Any individual who has an open credit account from a lender reporting to the credit bureaus (mortgage, auto loan, credit card, etc.), or who previously had an account that closed within the last 7 years has a credit history.

The sample is constructed using Social Security numbers ending in an arbitrarily chosen final two digits. This produces a random sample because the Social Security Administration assigns the last 4 digits of Social Security numbers sequentially, regardless of location. The panel tracks individuals over time, and allows people to enter and exit at the same rate as the target population, ensuring that the sample remains representative. This sampling method closely follows that of the Federal Reserve Bank of New York Consumer Credit Panel (see Lee and van der Klaauw (2010)). The data set is based on credit files as of Dec. 31st each year, and includes annual observations for approximately 2.3 million people per year from

2009 to 2015. The tests focus on 2010–2015 in order to use lagged control variables.

The credit bureau data provide a complete credit history for each individual, including the person's credit score, total debt, debt by category (mortgage, auto, credit card, etc.), past due debt, new sources of credit opened, and "hard" credit inquiries. These credit inquiries occur when a borrower applies for credit, and the lender checks their credit report. The data also provide the person's age and the census tract they live in.

C. HMDA Mortgage Application Data

The Home Mortgage Disclosure Act requires nearly all mortgage lenders to report detailed information on the applications they receive, and whether they originate the loan. Only very small or exclusively rural lenders are exempt from HMDA reporting.[7] Therefore, the HMDA database covers at least 95% of all first-lien mortgages in MSAs (Avery et al. (2017)). The data include requested loan size, income, race, and ethnicity as well as the purpose of the loan (purchase, refinancing, and home improvement), any co-applicants, and the loan's priority (first or second lien). The census tract of the property is also reported.

To construct the sample of mortgage applications for this paper, I merge lenders in the HMDA data to banks in the Call Report data based on federal agency identifiers common to both databases, and based on names for the remaining unmatched banks as in Loutskina and Strahan (2009). I select all mortgage applications received by commercial banks that are required to report HMDA data. I then exclude applications that the lender did not make a decision on due to the application being incomplete or withdrawn. Next, I require the application to be for a conventional mortgage (excludes applications related to programs run by the Federal Housing Administration, Veterans Administration, Farm Service Agency, or Rural Housing Service). I limit the sample to first-lien home purchase mortgage applications that are for loan amounts below the Government Sponsored Entities' securitization limits (excludes "jumbo" loans). Finally, I require the property to be located within 20 miles of the bank's nearest branch, and within an MSA, where HMDA data are the most comprehensive.[8] This process results in a sample of just over 3.6 million conventional mortgage applications between 2010 and 2015.

D. Intergenerational Mobility Data

I use county-level data on intergenerational mobility published by Chetty et al. (2014). The authors obtained access to records from the Social Security

[7]Depository institutions must report to the HMDA database if they have at least one branch or office in a metropolitan statistical area (MSA), have at least $44 million in assets (2016 threshold), and originated at least one mortgage in the previous year. Non-depository institutions with assets over $10 million must report if their mortgage originations total at least $25 million (or represent 10% of their loans), and they receive at least 5 mortgage applications in MSAs.

[8]None of the empirical results using these data are sensitive to these requirements. The results are similar if the MSA restriction is relaxed, or if applications further from branches are included, or if the distance variable is capped at a chosen maximum distance (e.g., 20, 30, 40, or 50 miles).

Administration and Internal Revenue Service and were able to link children born from 1980 to 1982 to their parents based on dependents on tax returns. Parental household income is measured as the average combined income of parent(s) from 1996 to 2000 (i.e., when the child is 15–19 years old), and the children's income is measured at age 26 (i.e., 2006–2008). The authors' sample includes 9.9 million children matched to their parents.

Based on these administrative data, Chetty et al. (2014) construct county-level intergenerational mobility statistics. Specifically, the authors provide estimates of the slope of child income on parent income for the people in a given county. This parent–child income slope is the coefficient from a rank–rank regression of child income distribution centile on parent income distribution centile (using the national income distribution). The authors also report transition matrices that describe the probabilities a child ends up in each quintile of the income distribution, based on which quintile their parents were in. The two measures of mobility I use are the parent–child income slope, and the probability that a child with parents in the bottom 40% of the income distribution moves out of this bottom 40% as an adult. I also examine the sensitivity of children's college attendance to their parent's income using data provided by Chetty et al. (2014).

IV. Large Bank Market Share and Households' Access to Credit

A. Baseline OLS Results

In this section, I test whether the size of local banks affects households' access to credit using data from a major credit bureau. I regress individual-level measures of credit approval on LARGE_BANK_MARKET_SHARE, which is defined as the fraction of bank branches within 10 miles of the household that are owned by banks with assets over $1 billion in 2010 dollars.[9] I measure credit approval based on the success rate of households' annual credit shopping attempts. Specifically, I select all person-years in the credit bureau data in which someone applies for credit based on the "hard" credit inquiry that appears on their file when a lender checks their credit score. I then construct an indicator, CREDIT_APPROVAL, which equals 1 when the person successfully opens a new source of credit during the year. I focus on overall CREDIT_APPROVAL across all types of credit, excluding credit cards, since credit card lending is dominated by a few national lenders and is unlikely to depend on local bank branches. But, the results also hold within various types of credit, and are robust to including credit card applications. Several recent papers that use credit bureau data construct and validate similar measures of credit access (e.g., Bhutta and Keys (2016), Akey, Dobridge, Heimer, and Lewellen (2018),

[9]The $1 billion cutoff follows prior studies in this literature (e.g., Berger and Black (2011), Berger et al. (2017)), but the results also hold using alternate definitions such as a $10 billion cutoff (see Table IA5 in the Supplementary Material). I compute distances between households and bank branches based on longitude and latitude coordinates using the Haversine formula. Households' coordinates are defined as the centroid of their census tract, and bank branch coordinates are available from the FDIC.

Brown, Cookson, and Heimer (2019), Akey, Heimer, and Lewellen (2021), and Butler, Mayer, and Weston (2023)).

Only those who apply for credit will be included in the tests, but this helps isolate credit supply. In fact, Heckman estimation methods produce larger estimates than the OLS results below (see Table IA3 in the Supplementary Material). Importantly, this annual measure of credit access is conservative in nature, because if one lender denies an applicant, but another lender steps in and makes the loan, the episode counts as a successful credit shopping attempt. This allows borrowers to shop around and end up at the type of bank most willing to lend to them (either knowingly or by trial and error shopping), which will work against finding that the composition of local banks affects credit access.

To test whether LARGE_BANK_MARKET_SHARE has a heterogeneous effect on households of high versus low economic status, I interact it with indicators for the person having a low income, subprime credit score, or limited credit history. LOW_INCOME indicates the person's ESTIMATED_INCOME from the credit bureau's proprietary model at the end of the prior year was below the median. This model is developed by the credit bureau based on a large sample of individuals' reported incomes on IRS tax returns and all of the individual attributes the credit bureau has on file, and it is re-verified annually. SUBPRIME indicates the person's CREDIT_SCORE at the end of the prior year was less than or equal to 660 (approximately 43% of the sample is subprime).[10] LIMITED_HISTORY indicates the person had below the median number of open credit lines at the end of the prior year (two or fewer).

The regressions of CREDIT_APPROVAL on LARGE_BANK_MARKET_SHARE also include individual characteristics as of the end of the prior year, census tract characteristics, county-level variables, and state-by-year fixed effects. Panel A of Table 1 presents summary statistics for the sample of approximately 2.3 million people per year from 2010 to 2015. Panel B summarizes how often individuals apply for various types of credit.

To allow for nonlinearities, I control for several of the individual characteristics using indicators based on binned values. The bins are based on 10 point intervals for CREDIT_SCORE, 5% ventiles for ESTIMATED_INCOME, and on each unique value for NUMBER_OF_CREDIT_LINES and AGE. These bin indicators for CREDIT_SCORE, ESTIMATED_INCOME, and NUMBER_OF_CREDIT_LINES absorb the direct effects of SUBPRIME, LOW_INCOME, and LIMITED_HISTORY when interacting these variables with LARGE_BANK_MARKET_SHARE. The remaining individual characteristics control for the amount of total debt and delinquent debt the person has. The census tract variables describe the local population where the person lives, and proxy for non-financial personal characteristics. The county-level variables control for local economic conditions. Finally, the state-by-year fixed effects are important because they control for differences in state policies that might affect credit supply, such as foreclosure or debt collection laws.

[10]The credit score I use throughout the paper is the Vantage Score. The 3 major consumer credit bureaus developed Vantage Score to rival FICO scores, and it is the second most popular credit score. Vantage Score has the same score range as FICO, and is very similar, which led FICO to sue (unsuccessfully) the credit bureaus for producing such a similar product.

TABLE 1

Summary Statistics

Table 1 presents summary statistics for the 1% national sample of individual credit bureau records used in the first set of empirical tests. The sample includes approximately 2.3 million annual observations each year from 2010 to 2015. Panel A summarizes the credit bureau variables as well as those describing the local banks, populations, and economies where individuals live. Columns 1–5 describe the full sample, and columns 6–8 describe the sample used in the instrumental variables approach based on state borders where states have large differences in policies toward interstate bank branching. Panel B presents statistics describing how often borrowers apply for certain types of credit. These application rates are reported for the full sample, for applicants with prime credit scores (CREDIT_SCORE > 660), and for the remaining applicants with subprime scores.

Panel A. Summary Statistics

	Full Sample (N = 13,833,955)					State Borders IV Sample (N = 2,582,708)		
	Mean	Std. Dev.	P10	P50	P90	Mean	Std. Dev.	Norm. Diff.
Credit Bureau Variables								
CREDIT_APPROVAL	0.6840	0.4649	0	1	1	0.7108	0.4534	0.0413
CREDIT_SCORE$_{t-1}$	674	111	516	678	813	675	112	0.0056
ESTIMATED_INCOME$_{t-1}$	45,561	25,699	23,000	39,000	75,000	43,819	22,970	−0.0505
NUMBER_OF_CREDIT_LINES$_{t-1}$	4.14	4.22	0.00	3.00	10.00	4.04	4.16	−0.0167
AGE	50	19	26	49	77	51	19	0.0084
log(TOTAL_DEBT$_{t-1}$)	7.05	4.84	0.00	8.76	12.29	6.97	4.85	−0.0105
TOTAL_DEBT$_{t-1}$	65,430	122,127	0	6380	218,206	59,726	108,684	−0.0349
log(PAST_DUE_DEBT$_{t-1}$)	2.52	3.62	0.00	0.00	8.33	2.52	3.58	−0.0011
PAST_DUE_DEBT$_{t-1}$	1591	4796	0	0	4152	1450	4418	−0.0216
HAVE_DELINQUENT_DEBT$_{t-1}$	0.2020	0.4015	0	0	1	0.1945	0.3958	−0.0132
Local Banks								
LARGE_BANK_MARKET_SHARE	0.7698	0.2151	0.4667	0.8495	0.9498	0.6965	0.2299	−0.2329
HHI_OF_LOCAL_BANK_BRANCHES	0.1356	0.1256	0.0656	0.1000	0.2222	0.1349	0.1268	−0.0044
Census Tract Characteristics								
POVERTY_(18–64)	0.1304	0.0976	0.0310	0.1050	0.2670	0.1328	0.0974	0.0173
log(POPULATION_DENSITY)	7.17	1.94	4.21	7.65	9.21	6.83	1.90	−0.1254
MINORITY_POPULATION_SHARE	0.3475	0.2894	0.0470	0.2530	0.8430	0.2553	0.2479	−0.2420
HOUSEHOLD_SIZE	2.66	0.47	2.13	2.60	3.25	2.57	0.36	−0.1420
HIGH_SCHOOL_DIPLOMA	0.8632	0.1069	0.7170	0.8920	0.9690	0.8648	0.0943	0.0111
EMPLOYED_BY_GOVERNMENT	0.1476	0.0668	0.0720	0.1370	0.2370	0.1386	0.0616	−0.0992
County Characteristics								
UNEMPLOYMENT_RATE	0.0761	0.0245	0.0467	0.0733	0.1090	0.0777	0.0229	0.0484
PIPC_GROWTH	0.0329	0.0267	−0.0011	0.0341	0.0633	0.0316	0.0264	−0.0338
ESTABLISHMENT_GROWTH	−0.0011	0.0184	−0.0247	−0.0002	0.0212	−0.0045	0.0177	−0.1318
HOUSE_PRICE_GROWTH	0.0092	0.0559	−0.0545	0.0037	0.0871	−0.0041	0.0414	−0.1924

Panel B. Credit Application Rates

Credit Application Type	Fraction of Person-Years with Credit Applications		
	Full Sample	Prime Borrowers	Subprime Borrowers
All types	0.5432	0.5279	0.5639
Mortgage	0.1349	0.1557	0.1068
Auto	0.1415	0.1311	0.1555
Credit card	0.2738	0.2729	0.275
All non-credit card	0.4563	0.4389	0.4796

Table 2 presents the baseline OLS results. The regression in column 1 shows that a standard deviation increase in LARGE_BANK_MARKET_SHARE leads to a 0.40-percentage-point decrease in CREDIT_APPROVAL across all borrowers, compared to the mean CREDIT_APPROVAL of 68.4%. The tests in columns 2–4 show that this result in column 1 is driven by a much larger reduction in credit access for individuals of low economic status, whereas borrowers of high economic status are relatively unaffected. For instance, column 3 shows that a standard deviation increase in LARGE_BANK_MARKET_SHARE leads to a 0.09-percentage-point decrease in CREDIT_APPROVAL for borrowers with prime credit scores, whereas

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

TABLE 2

Large Bank Market Share and Household Credit Access: Baseline OLS Results

Table 2 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 sample includes all person-years in the credit bureau data set 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 test in column 1 estimates the effect of LARGE_BANK_MARKET_SHARE for all applicants. Columns 2–4 interact LARGE_BANK_MARKET_SHARE with indicators for the applicant having a low income, low credit score, or limited credit history, respectively. LOW_INCOME indicates the applicant's estimated income from the credit bureau's proprietary model is below the median. SUBPRIME indicates the applicant has a CREDIT_SCORE ≤ 660 (43% of people have subprime scores). LIMITED_HISTORY indicates the applicant had below the median number of open credit lines at the end of the prior year (2 or fewer). The base terms for the interaction between these three variables and LARGE_BANK_MARKET_SHARE are omitted because they are direct linear combinations of the bin indicators I use to control for the direct effect (i.e., bin indicators for CREDIT_SCORE, ESTIMATED_INCOME, and NUMBER_OF_CREDIT_LINES). All continuous explanatory variables are standardized to have a mean of 0 and a standard deviation of 1. Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	1	2	3	4
LARGE_BANK_MARKET_SHARE	−0.402***	−0.103***	−0.0948***	−0.122***
	(0.0254)	(0.0287)	(0.0278)	(0.0270)
LARGE_BANK_MARKET_SHARE × LOW_INCOME		−0.637***		
		(0.0340)		
LARGE_BANK_MARKET_SHARE × SUBPRIME			−0.669***	
			(0.0357)	
LARGE_BANK_MARKET_SHARE × LIMITED_HISTORY				−0.789***
				(0.0374)
Individual Characteristics				
CREDIT_SCORE_10-POINT_BIN_INDICATORS	Yes	Yes	Yes	Yes
ESTIMATED_INCOME_VENTILE_INDICATORS	Yes	Yes	Yes	Yes
NUMBER_OF_CREDIT_LINES_INDICATORS	Yes	Yes	Yes	Yes
AGE_INDICATORS	Yes	Yes	Yes	Yes
log(TOTAL_DEBT$_{t-1}$)	−1.724***	−1.742***	−1.724***	−1.721***
	(0.0485)	(0.0486)	(0.0485)	(0.0485)
log(PAST_DUE_DEBT$_{t-1}$)	−3.766***	−3.773***	−3.771***	−3.770***
	(0.0304)	(0.0304)	(0.0304)	(0.0304)
HAVE_DELINQUENT_DEBT$_{t-1}$	−3.975***	−3.983***	−3.982***	−3.972***
	(0.0541)	(0.0541)	(0.0541)	(0.0541)
Census Tract Characteristics				
HHI_OF_LOCAL_BANK_BRANCHES_(10MI)	−0.0863***	−0.0949***	−0.0934***	−0.0968***
	(0.0227)	(0.0227)	(0.0227)	(0.0227)
POVERTY_(18–64)	−0.0597**	−0.0554**	−0.0560**	−0.0527*
	(0.0277)	(0.0277)	(0.0277)	(0.0277)
log(POPULATION_DENSITY)	−1.110***	−1.123***	−1.128***	−1.126***
	(0.0300)	(0.0300)	(0.0300)	(0.0300)
MINORITY_POPULATION_SHARE	−1.133***	−1.114***	−1.108***	−1.115***
	(0.0323)	(0.0323)	(0.0324)	(0.0323)
HOUSEHOLD_SIZE	0.922***	0.920***	0.925***	0.922***
	(0.0241)	(0.0241)	(0.0241)	(0.0241)
HIGH_SCHOOL_DIPLOMA	−0.675***	−0.684***	−0.680***	−0.682***
	(0.0315)	(0.0315)	(0.0315)	(0.0315)
EMPLOYED_BY_GOVERNMENT	0.917***	0.915***	0.916***	0.916***
	(0.0209)	(0.0209)	(0.0209)	(0.0209)
County Characteristics				
UNEMPLOYMENT_RATE	0.412***	0.405***	0.407***	0.401***
	(0.0351)	(0.0351)	(0.0351)	(0.0351)
PIPC_GROWTH	−0.201***	−0.203***	−0.203***	−0.203***
	(0.0249)	(0.0248)	(0.0248)	(0.0249)
ESTABLISHMENT_GROWTH	−0.453***	−0.451***	−0.451***	−0.452***
	(0.0288)	(0.0288)	(0.0288)	(0.0288)
HOUSE_PRICE_GROWTH	0.727***	0.716***	0.725***	0.721***
	(0.0478)	(0.0478)	(0.0478)	(0.0478)
State × year FE	Yes	Yes	Yes	Yes
R^2	0.194	0.194	0.194	0.194
No. of obs.	6,202,881	6,202,881	6,202,881	6,202,881

it leads to a 0.76-percentage-point decrease for subprime borrowers, whose mean CREDIT_APPROVAL is 53%. A similar pattern emerges in the results in columns 2 and 4 when LOW_INCOME or LIMITED_HISTORY are used to define borrowers of low economic status.

These baseline findings are robust to a variety of alternate empirical approaches. Table IA4 in the Supplementary Material shows that the results are similar when the local banking market is defined based on the borrower's ZIP code rather than a 10-mile surrounding area, and when either branches or deposits are used to compute LARGE_BANK_MARKET_SHARE. I focus on the 10-mile surrounding area throughout the paper because it is more uniform and allows local banking markets to cross state lines, and this is necessary for the instrumental variables design. Table IA4 also shows that these 10-mile areas are roughly halfway between ZIP codes and counties in terms of the number of banks, and that LARGE_BANK_MARKET_SHARE is similar when measured at each geographic level. Finally, Table IA5 in the Supplementary Material implements two alternate measures of large banks' local presence: a measure that uses a $10 billion cutoff to delineate large versus small banks, and a measure that excludes the 4 largest U.S. banks (JP Morgan Chase, Bank of America, Wells Fargo, and Citibank) from the computation of LARGE_BANK_ MARKET_SHARE. In each case, the baseline results hold.

B. Identification Issues and Additional OLS Results

One potential concern with the baseline OLS results is that there is a large amount of variation across U.S. cities and neighborhoods in terms of both the local bank characteristics and the local population. The rich set of time-varying controls at the individual, census tract, and county level help control for much of this variation. However, if omitted geographic factors correlate with both LARGE_BANK_MARKET_SHARE and aspects of borrower creditworthiness not captured in credit scores, an omitted variables problem could arise. Therefore, the next tests implement census tract fixed effects that control for any persistent differences across neighborhoods.

Panel A of Table 3 presents results using census tract fixed effects. These tests use only the time-series variation in local banks that neighborhood residents have access to. In other words, the variation in LARGE_BANK_MARKET_SHARE over time that is driven by large and small banks entering and exiting local markets. The results in Panel A look similar to the baseline OLS results: LARGE_ BANK_MARKET_SHARE reduces credit access, especially for borrowers of low economic status.

A particularly salient case to study is the extensive margin (i.e., whether borrowers have access to *any* small banks). Panel B of Table 3 presents tests similar to those in Panel A, except replacing LARGE_BANK_MARKET_SHARE with NO_SMALL_BANKS, an indicator for the person living in a census tract that currently has no branches of small banks within 10 miles. The within-tract time-series variation in NO_SMALL_BANKS comes from when a neighborhood gains access to its first small bank, or when the last local small bank exits the market. The results in Panel B show that having no small banks has no effect on higher income

TABLE 3

Large Bank Market Share and Household Credit Access: Neighborhood Fixed Effects

Table 3 presents regressions of individuals' CREDIT_APPROVAL on measures of large banks' market share, and time-varying controls for individual, census tract, and county-level characteristics, as well as census tract and year fixed effects. The sample includes all person-years in the credit bureau data set from 2010 to 2015 where the person applies for credit. CREDIT_APPROVAL is an indicator for the individual successfully opening a new credit line. In Panel A, the test in column 1 estimates the effect of LARGE_BANK_MARKET_SHARE for all applicants. Columns 2, 3, and 4 interact LARGE_BANK_MARKET_SHARE with indicators for the applicant having a low income, low credit score, or limited credit history, respectively. The base terms for the interaction between these three variables and LARGE_BANK_MARKET_SHARE are omitted because they are direct linear combinations of the bin indicators I use to control for the direct effect (i.e., bin indicators for CREDIT_SCORE, ESTIMATED_INCOME, and NUMBER_OF_CREDIT_LINES). Panel B presents similar tests, where LARGE_BANK_MARKET_SHARE is replaced with NO_SMALL_BANKS, an indicator for the person living in a census tract that currently has no branches of small banks (less than $1 billion in assets) within 10 miles. All continuous explanatory variables are standardized to have a mean of 0 and a standard deviation of 1. Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Census Tract Fixed Effects

	1	2	3	4
LARGE_BANK_MARKET_SHARE	−0.397***	−0.214***	−0.162**	−0.179**
	(0.0736)	(0.0750)	(0.0750)	(0.0744)
LARGE_BANK_MARKET_SHARE × LOW_INCOME		−0.428***		
		(0.0352)		
LARGE_BANK_MARKET_SHARE × SUBPRIME			−0.509***	
			(0.0370)	
LARGE_BANK_MARKET_SHARE × LIMITED_HISTORY				−0.632***
				(0.0385)
Individual, census tract, and county controls	Yes	Yes	Yes	Yes
Census tract FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.208	0.208	0.208	0.208
No. of obs.	6,202,672	6,202,672	6,202,672	6,202,672

Panel B. Extensive Margin

	1	2	3	4
NO_SMALL_BANKS	−0.184	0.0827	0.0952	0.0258
	(0.211)	(0.227)	(0.227)	(0.220)
NO_SMALL_BANKS × LOW_INCOME		−0.566***		
		(0.204)		
NO_SMALL_BANKS × SUBPRIME			−0.593***	
			(0.206)	
NO_SMALL_BANKS × LIMITED_HISTORY				−0.635***
				(0.222)
Individual, census tract, and county controls	Yes	Yes	Yes	Yes
Census tract FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
R^2	0.208	0.208	0.208	0.208
No. of obs.	6,202,672	6,202,672	6,202,672	6,202,672

households, but it significantly reduces credit access for households of lower economic status.

By isolating the within-neighborhood time-series variation in local bank characteristics, the tests in Table 3 help rule out alternative explanations based on persistent differences across neighborhoods. Yet, potential biases may still arise due to large banks choosing when to build/buy branches in a neighborhood. Importantly, if large banks systematically put branches where they want to lend to local households, potentially based on expected local economic conditions or household characteristics that are difficult to control for, then we should expect LARGE_BANK_MARKET_SHARE to be positively correlated with the unobserved component of borrower creditworthiness. Clearly, this would bias the OLS estimate of the effect of LARGE_BANK_MARKET_SHARE on CREDIT_APPROVAL upward (i.e., *against* the baseline OLS findings).

On the other hand, large banks may consider other factors when making branch location decisions, such as where they expect to attract deposits or lend to local businesses. Table IA6 in the Supplementary Material replicates the baseline results in various subsamples where these factors could be more or less prominent, such as urban/rural areas and high/low inequality areas.[11] The results show that the negative effect of LARGE_BANK_MARKET_SHARE on CREDIT_APPROVAL is not restricted to, or driven by, a certain type of location, helping to raise the bar for alternative explanations. The next section (IV.C) provides further support for a causal interpretation by exploiting exogenous variation in LARGE_BANK_MARKET_SHARE along state borders with contrasting policies toward bank branching.

C. Instrumental Variables Approach

I use an instrumental variables approach to isolate exogenous variation in LARGE_BANK_MARKET_SHARE and avoid any omitted variables bias resulting from large banks choosing where to locate their branches. The approach exploits the differences in state policies toward interstate bank branching that emerged following the IBBEA as discussed in Section II. These policies directly affect the ability of out-of-state banks (e.g., national banks) to enter local markets through building new branches or purchasing existing ones. I follow Rice and Strahan (2010) and use an index that describes the number of regulatory restrictions that out-of-state banks face when they consider establishing a branch in a state. The index ranges from 0 to 4 and increases by 1 if the state restricts the ability of out-of-state banks to build de novo branches, or to purchase individual branches of an existing bank. The index also increases if the state requires target banks in an interstate merger to have less than a 30% share of the state's deposits, or to be at least 3 years old.

Based on the index, I identify 36 state borders where one state has strong barriers to out-of-state bank branching (3 or 4 barriers), and the other state is open to out-of-state bank entry (0 or 1 barrier).[12] I find that these regulatory barriers affect LARGE_BANK_MARKET_SHARE: branches in the states with strong barriers are owned by smaller banks. To exploit this variation, I select everyone in the credit bureau data living within 50 miles of these borders and use their location relative to the border to instrument for LARGE_BANK_MARKET_SHARE. Figure 2 presents a map of the continental United States with the census tracts in these border areas highlighted.

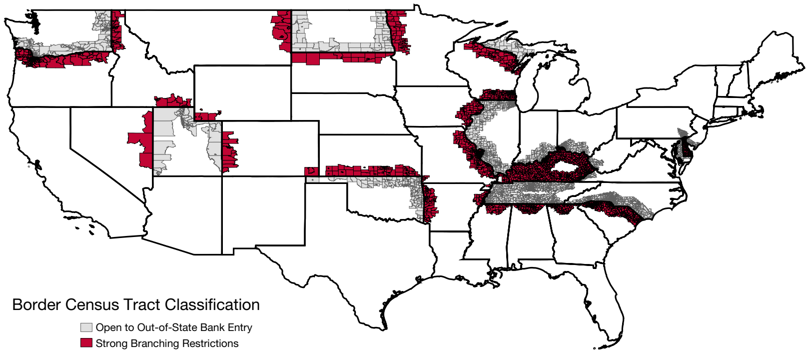
[11]For brevity, these tests focus on the main effect of LARGE_BANK_MARKET_SHARE and its interaction with SUBPRIME, but the results look similar using interactions with LOW_INCOME or LIMITED_HISTORY.

[12]See Table IA2 in the Supplementary Material for a list of these state borders. I assign each state its value of the index as of 2010, when the Dodd-Frank Act eliminated states' ability to restrict de novo branching. This approach ensures that states that prevented de novo branching from 1994 to 2010 are classified as having been more difficult for out-of-state banks to enter than states that allowed de novo branching during this period.

FIGURE 2

State Borders with a Large Contrast in Branching Restrictions

Figure 2 shows the state borders where there is a large contrast in the two states' interstate bank branching policies. I use the index of branching restrictions developed in Rice and Strahan (2010), which ranges from 0 to 4, to define states with 3 or 4 restrictions as having strong restrictions, and to define states with 0 or 1 restriction as being open to out of state bank entry. This map shows the census tracts within 50 miles of the 36 state borders where one state has strong restrictions and the other state is open to out of state bank entry (see Table IA2 in the Supplementary Material for a list of these state borders).



The instrumental variable I use, POSITION_RELATIVE_TO_BORDER, ranges from −50 in the interior of states with strong regulatory barriers, to 50 in the interior of states that are open to out-of-state bank entry. POSITION_RELATIVE_TO_BORDER has a positive effect on LARGE_BANK_ MARKET_SHARE because banks in the states that are open to out-of-state bank entry are larger. Graph A in Figure 3 shows the relationship between a census tract's POSITION_RELATIVE_TO_BORDER and the residual fraction of the bank branches in the tract that are owned by large banks. These residuals are from a census tract-level regression of the large bank share in the tract on tract characteristics and year fixed effects. The graph shows that conditional on census tract characteristics, large banks indeed own a higher percentage of branches in states that are open to out-of-state bank branching.

Graphs B and C in Figure 3 show how CREDIT_APPROVAL varies across state borders for borrowers with prime and subprime credit scores, respectively. These figures plot the residual CREDIT_APPROVAL from an individual-level regression, against the person's POSITION_RELATIVE_TO_ BORDER. The individual-level regression includes all of the controls from the previous tests in Tables 2 and 3 except LARGE_BANK_MARKET_SHARE and the state-by-year fixed effects. The figures suggest that both prime and subprime borrowers experience greater credit access when local banks are small, but the effect appears to be larger for subprime borrowers. The ensuing instrumental variables regressions augment and formalize this test.

Panel A of Table 4 presents the first-stage regressions for the instrumental variables approach. Column 1 shows the results when LARGE_BANK_ MARKET_SHARE (LBMS) is regressed on POSITION_RELATIVE_TO_ BORDER and individual, census tract, and county-level controls, as well as state-by-year fixed effects. Because this paper tests whether the effect of LARGE_BANK_MARKET_SHARE is different for households of low economic

FIGURE 3

Bank Size and Household Credit Access Across State Borders with a Large Contrast in Interstate Bank Branching Policies

Figure 3 shows how the size of banks and household credit access change around state borders where the two states have a stark contrast in interstate bank branching policies. I use the index of branching restrictions developed in Rice and Strahan (2010), which ranges from 0 to 4, to define states with 3 or 4 restrictions as having strong restrictions, and to define states with 0 or 1 restriction as being open to out of state bank entry. Graph A shows the residual share of branches owned by large banks (assets greater than 1 billion in 2010 dollars) in census tracts based on the tract's position relative to the border (measured in miles). These residuals are from a census tract-level regression of the large bank share on tract characteristics and year fixed effects. Graph B shows how residual credit approval varies across the relevant state borders for prime borrowers. Residual credit approval is obtained from an individual-level regression of CREDIT_APPROVAL on the individual, census tract, and county-level controls (see Table 2). Graph C shows how residual credit approval varies across these borders for subprime borrowers.

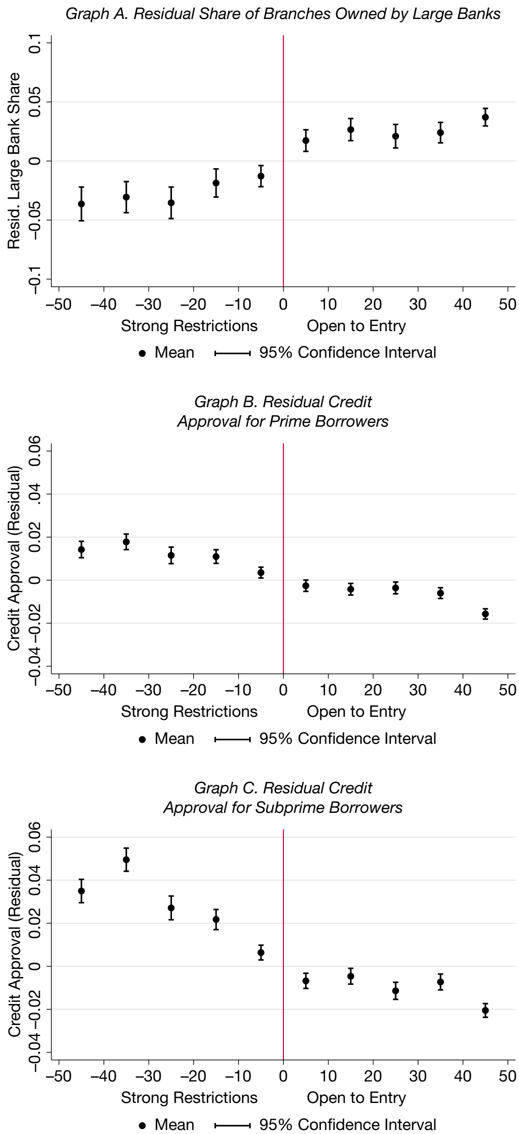


TABLE 4

IV/2SLS Analysis of the Effect of Large Bank Market Share on Household Credit Access

Table 4 presents the IV/2SLS analysis of the effect of LARGE_BANK_MARKET_SHARE (LBMS) on households' credit access. Panel A presents the first-stage regressions. The dependent variable in column 1 is LBMS, which is the fraction of bank branches within 10 miles of the household that are owned by banks with greater than $1 billion in assets. The dependent variables in columns 2–4 are the interactions between LBMS and the indicators of low economic status (LOW_INCOME, SUBPRIME, and LIMITED_HISTORY). The sample includes all person-years in the credit bureau data from 2010 to 2015 where the person applies for credit and 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 in the Supplementary Material for a list of these state borders). The instrumental variables are POSITION_RELATIVE_TO_BORDER and its interaction with the indicators of low economic status. POSITION_RELATIVE_TO_BORDER ranges from −50 to 50, with −50 representing census tracts that are 50 miles toward the interior of the state with strong branching restrictions, and positive values representing tracts toward the interior of the state that is open to out of state bank entry. The regressions also control for personal characteristics from the credit bureau data, and census tract and county-level characteristics, as well as state-by-year fixed effects. Panel B presents the OLS and IV estimates of the effect of LARGE_BANK_MARKET_SHARE on CREDIT_APPROVAL, which is an indicator for whether a person applying for credit is successful in opening a new credit line. As shown in Panel A, we instrument for LBMS and its interactions with indicators of low economic status with POSITION_RELATIVE_TO_BORDER and its interactions with the economic status indicators. 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). All continuous explanatory variables are standardized to have a mean of 0 and a standard deviation of 1 (except POSITION_RELATIVE_TO_BORDER, which is in miles). Coefficients are reported in percentage point units, and the standard errors are clustered by census tract-year. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. First-Stage Regressions

	LBMS	LBMS × LOW_INCOME	LBMS × SUBPRIME	LBMS × LIMITED_HISTORY
	1	2	3	4
POSITION_RELATIVE_TO_BORDER	0.149***	0.141***	0.0942***	0.0817***
	(0.0136)	(0.0284)	(0.0268)	(0.0214)
POSITION_RELATIVE_TO_BORDER × LOW_INCOME		0.355***		
		(0.0453)		
POSITION_RELATIVE_TO_BORDER × SUBPRIME			0.371***	
			(0.0446)	
POSITION_RELATIVE_TO_BORDER × LIMITED_HISTORY				0.348***
				(0.0482)
Individual, census tract, and county controls	Yes	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes	Yes
R^2	0.570	0.307	0.285	0.239
No. of obs.	1,125,673	1,125,673	1,125,673	1,125,673

Panel B. OLS and IV Estimates of the Effect of LBMS on Credit Approval

	OLS				IV			
	1	2	3	4	5	6	7	8
LBMS	−0.506***	−0.0869	0.122*	−0.0554	−1.015**	0.666	1.079**	0.717
	(0.0593)	(0.0672)	(0.0643)	(0.0626)	(0.499)	(0.510)	(0.500)	(0.502)
LBMS × LOW_INCOME		−0.872***				−3.535***		
		(0.0774)				(0.414)		
LBMS × SUBPRIME			−1.432***				−4.754***	
			(0.0830)				(0.444)	
LBMS × LIMITED_HISTORY				−1.274***				−5.047***
				(0.0850)				(0.462)
Individual, census tract, and county controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
State × year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.196	0.196	0.196	0.196	–	–	–	–
No. of obs.	1,125,673	1,125,673	1,125,673	1,125,673	1,125,673	1,125,673	1,125,673	1,125,673
First-stage *F*-stat	–	–	–	–	838.7	420.2	421.4	420.1

status, I also instrument for the interaction between LARGE_BANK_MARKET_SHARE and indicators of low economic status (LOW_INCOME, SUBPRIME, and LIMITED_HISTORY). Columns 2–4 show the first-stage regressions for these interaction terms, which are instrumented for with the interaction

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

between POSITION_RELATIVE_TO_BORDER and the corresponding economic status indicator. The results in columns 1–4 show that in all cases, POSITION_RELATIVE_TO_BORDER and its interactions, are strong predictors of LARGE_BANK_MARKET_SHARE and its interactions. The instruments' power comes from the fact that the regulatory barriers make it more costly for large out-of-state banks to enter local markets in one state. The identifying assumption necessary to satisfy the exclusion restriction is that, for two borrowers in the same state during the same year, controlling for credit scores and individual, census tract, and county-level characteristics, their distance to the state border affects CREDIT_APPROVAL only through its effect on LARGE_BANK_MARKET_SHARE.

Panel B of Table 4 presents the main instrumental variables regressions and their OLS counterparts. In the OLS regressions in columns 1–4, the same pattern in the coefficients of interest emerges as in the baseline OLS results: when estimated across all borrowers, LARGE_BANK_MARKET_SHARE has a negative effect on CREDIT_APPROVAL, but this effect is driven almost entirely by the effect on borrowers of low economic status. This pattern is even more striking in the instrumental variables tests in columns 5–8. For instance, the results in column 7 show that a standard deviation increase in LARGE_BANK_MARKET_SHARE actually increases prime borrowers' CREDIT_APPROVAL by 1.08 percentage points compared to a mean of 80.9%. In contrast, for subprime borrowers a standard deviation increase in LARGE_BANK_MARKET_SHARE reduces CREDIT_APPROVAL by 3.67 percentage points (1.08–4.75), compared to their mean CREDIT_APPROVAL of 53.0%.

Additional tests in Table IA7 in the Supplementary Material show that higher LARGE_BANK_MARKET_SHARE leads to lower credit approval rates for first-time mortgage applicants, to local households carrying a larger share of their debt on credit cards, and to increased borrowing from retailers. The lower mortgage approval rates and increased use of higher-cost debt are most pronounced among borrowers of low economic status, consistent with these households facing credit rationing from banks. Overall, the results in this section show that having large local banks reduces credit access for borrowers of low economic status, whereas borrowers of high economic status continue to receive credit and may even experience increased credit access. These patterns hold when measuring economic status based on income, credit score, or depth of credit history, and using various empirical approaches and designs.

V. Bank Size, Soft Information, and Lending to Households

A. Bank Size and Mortgage Approval

In this section, I test whether the relationship between local bank size and household credit constraints is driven by differences in soft information production at small versus large banks. Several studies show that soft information matters when lending to consumers. Iyer et al. (2016) show that non-experts on a peer-to-peer lending platform are able to outperform credit scores in predicting default using pictures of the prospective borrowers, their requested interest rate, and short pieces

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

of text explaining the reason for borrowing. Agarwal et al. (2011) use mortgage application data from a single bank to show that loan officers collect soft information from their interactions with borrowers as they handle an application. Overall, research in this area shows that the human component of the lending process can add value by incorporating soft information on factors such as the applicant's circumstances, motivation for taking out the loan, or even personal character or perseverance (Cornaggia, Cornaggia, and Xia (2021)), which are not directly captured in standard hard information, but which influence loan repayment.

Stein (2002) predicts that small banks' decentralized organizational structures put more authority in the hands of local loan officers and branch managers, and that this facilitates lending based on soft information. Consistent with this theory, prior studies show that small banks play an important role in small business lending (e.g., Berger et al. (2017)). However, it is unclear from these studies whether small banks play a special role in consumer credit markets, or in particular for low-income households whose access to credit may depend on soft information. Therefore, I test whether small banks incorporate more soft information when lending to households, and whether this information is particularly important for low-income households. If both are true, it could shed light on the mechanism behind the results in Section IV showing that low-income households face reduced access to credit when local banks are large.

To evaluate the extent to which banks use soft information, I examine how much credit approval rates decrease with borrower–lender distance. This approach follows an extensive literature in finance that interprets soft information collection as the main channel through which borrower–lender distance affects credit provision (e.g., Petersen and Rajan (2002), DeYoung et al. (2008), and Agarwal and Hauswald (2010)).

The tests in this section use data on mortgage applications from the Home Mortgage Disclosure Act. The HMDA data include the identity of the lender receiving the application, their decision to approve or deny the loan, and detailed characteristics of the applicant and the loan they requested. Following the process described in Section III, I construct a sample of just over 3.6 million conventional mortgage applications received by commercial banks between 2010 and 2015. Table 5 provides summary statistics, and shows that small banks (with assets less than $1 billion) receive 21% of the applications, that the average distance from the property to the bank's nearest branch is 3.8 miles, and that the average loan amount requested is $201,000.

To test whether soft information plays a larger role in lending decisions at small banks, I regress an indicator for the mortgage application being approved on the distance to the bank's nearest branch, its interaction with SMALL_BANK, and control variables. The controls include borrower, census tract, and bank characteristics, as well as county-by-year fixed effects. The fixed effects capture across-county variation in borrower creditworthiness, local housing market conditions, and state policies that affect mortgage credit availability (e.g., foreclosure laws). The applicant characteristics work to control for variation in the creditworthiness of applicants within the same county-year applying to small versus large banks. These characteristics include the applicant's income, the loan amount, the loan to income ratio, and an indicator for joint applications (multiple applicants). I also include

TABLE 5

Summary Statistics for HMDA Mortgage Applications

Table 5 presents summary statistics describing the sample of mortgage applications from the Home Mortgage Disclosure Act database. I collect all applications received by commercial banks for conventional mortgages (excludes applications related to programs run by the Federal Housing Administration, Veterans Administration, Farm Service Agency, or Rural Housing Service). I limit the sample to first-lien home purchase mortgage applications that are for loan amounts below the Government Sponsored Entities' securitization limits (excludes "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 sample includes just over 3.6 million mortgage applications between 2010 and 2015.

	Mean	Std. Dev.	P10	P50	P90
Main Variables of Interest					
MORTGAGE_APPROVAL	0.854	0.353	0	1	1
DISTANCE_TO_BRANCH	3.790	4.041	0.603	2.186	9.622
SMALL_BANK	0.207	0.405	0	0	1
LOW_INCOME	0.243	0.429	0	0	1
Applicant and Loan Characteristics					
INCOME	110,455	89,239	36,000	86,000	206,000
LOAN_AMOUNT	201,009	126,518	60,000	172,000	392,000
LOAN_TO_INCOME_RATIO	2.238	1.237	0.723	2.093	3.913
JOINT_APPLICATION	0.489	0.500	0	0	1
BLACK	0.035	0.184	0	0	0
HISPANIC	0.071	0.257	0	0	0
Census Tract Ratios and Averages					
INCOME/TRACT_INCOME	0.977	0.627	0.380	0.831	1.718
LOAN_TO_INCOME/TRACT_LOAN_TO_INCOME	1.012	0.512	0.381	0.965	1.677
LOAN_AMOUNT/TRACT_LOAN_AMOUNT	1.014	0.430	0.487	0.980	1.567
AVERAGE_CREDIT_SCOREt-1	689	39	633	693	736
Bank Characteristics					
CAPITAL_RATIO	0.105	0.033	0.076	0.108	0.134
REAL_ESTATE_LOANS_RATIO	0.415	0.164	0.234	0.406	0.647
PROFITABILITY	0.008	0.008	0.002	0.010	0.014

indicators for one or more of the applicants being a minority, given the literature on mortgage discrimination (e.g., Munnell, Tootell, Browne, and McEneaney (1996)).

A disadvantage of the HMDA data is that they do not include applicant credit scores. The credit bureau data also cannot be directly linked to all HMDA applications (although in the next section (V.B), I match the data sets for *originated* loans). However, I am able to reduce these concerns by controlling for the average credit score of residents in the census tract in which the applicant is trying to purchase a home. I also control for the ratio of the applicant's income, loan to income ratio, and loan amount, to the average of all applicants within the census tract that year. These variables are designed to capture whether the applicant is more or less creditworthy than the typical applicant in the census tract.

Table 6 presents the results. Column 1 shows that for the full sample of borrowers and banks, each additional mile between the parties reduces the chances of an application being approved by 10 basis points. Column 2 shows that the effect of distance on mortgage approval is over twice as large at small banks compared to large banks. Column 3 shows that the effect of distance is 3 times larger for low-income applicants than high-income applicants. Columns 4 and 5 split the sample based on whether the applicant has a low or high income and show that in each case distance matters more at small banks. The positive estimated effect of SMALL_BANK in these tests also demonstrates that small banks approve a higher percentage of mortgage applications overall than large banks.[13]

[13]Robustness tests in Panel B of Table IA5 in the Supplementary Material show that similar results hold when small banks are defined as those with assets under $10 billion (rather than $1 billion), and

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

TABLE 6

The Effect of Borrower–Lender Distance on Mortgage Approval at Small Versus Large Banks

Table 6 presents regressions of an indicator for a mortgage application being approved on the distance from the property to the bank's nearest branch, and this distance interacted with an indicator for the bank being small (assets less than 1 billion in 2010 dollars) or the borrower having a low income (below the median U.S. household income), as well as control variables. Columns 1–3 present the results for the full sample, which includes all first-lien home purchase mortgage applications in the Home Mortgage Disclosure Act database that were received by commercial banks from 2010 to 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). Columns 4 and 5 present the results for the subsample of low-income and high-income applicants, respectively. All specifications include county-by-year fixed effects, and the continuous explanatory variables are standardized to have a mean of 0 and a standard deviation of 1, except for DISTANCE_TO_BRANCH, which is in miles. Coefficients are reported in percentage point units, and the standard errors are clustered by county-year. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample			Low Income	High Income
	1	2	3	4	5
DISTANCE_TO_BRANCH	−0.0975***	−0.0762***	−0.0673***	−0.161***	−0.0417***
	(0.0102)	(0.0128)	(0.0100)	(0.0227)	(0.0123)
DISTANCE_TO_BRANCH × SMALL_BANK		−0.0793***		−0.0957***	−0.0715***
		(0.0165)		(0.0332)	(0.0157)
SMALL_BANK		0.912***		0.852***	1.079***
		(0.148)		(0.257)	(0.140)
DISTANCE_TO_BRANCH × LOW_INCOME			−0.126***		
			(0.0170)		
Applicant and Loan Characteristics					
INCOME_CENTILE_INDICATORS	Yes	Yes	Yes	Yes	Yes
LOAN_AMOUNT_CENTILE_INDICATORS	Yes	Yes	Yes	Yes	Yes
LOAN_TO_INCOME_RATIO	−3.523***	−3.534***	−3.531***	−7.956***	−1.671***
	(0.134)	(0.134)	(0.134)	(0.291)	(0.136)
JOINT_APPLICATION	−0.222***	−0.227***	−0.225***	−3.047***	0.797***
	(0.0520)	(0.0520)	(0.0520)	(0.120)	(0.0534)
BLACK	−8.014***	−8.011***	−8.021***	−9.406***	−7.302***
	(0.181)	(0.182)	(0.181)	(0.318)	(0.175)
HISPANIC	−4.091***	−4.085***	−4.101***	−4.761***	−3.628***
	(0.156)	(0.156)	(0.156)	(0.259)	(0.175)
Census Tract Ratios and Averages					
INCOME/TRACT_INCOME	0.110*	0.109	0.108	−1.862***	0.495***
	(0.0668)	(0.0668)	(0.0667)	(0.288)	(0.0610)
LOAN_TO_INCOME/TRACT_LOAN_TO_INCOME	−1.056***	−1.056***	−1.051***	−1.707***	−0.954***
	(0.0889)	(0.0888)	(0.0887)	(0.156)	(0.0874)
LOAN_AMOUNT/TRACT_LOAN_AMOUNT	−0.955***	−0.954***	−0.969***	−0.0340	−1.259***
	(0.0630)	(0.0629)	(0.0631)	(0.151)	(0.0613)
AVERAGE_CREDIT_SCOREt-1	1.118***	1.119***	1.121***	0.753***	1.272***
	(0.0491)	(0.0491)	(0.0490)	(0.0795)	(0.0550)
Bank Characteristics					
CAPITAL_RATIO	−0.498***	−0.490***	−0.499***	−0.0772	−0.615***
	(0.0715)	(0.0715)	(0.0715)	(0.104)	(0.0709)
REAL_ESTATE_LOANS_RATIO	3.965***	3.882***	3.963***	4.577***	3.653***
	(0.0669)	(0.0731)	(0.0670)	(0.105)	(0.0743)
PROFITABILITY	0.0799	0.0979*	0.0819	−0.0702	0.158***
	(0.0563)	(0.0566)	(0.0563)	(0.0876)	(0.0558)
County × year FE	Yes	Yes	Yes	Yes	Yes
R^2	0.082	0.082	0.082	0.119	0.053
No. of obs.	3,619,343	3,619,343	3,619,343	881,016	2,738,264

when the top 4 largest U.S. banks are excluded from the analysis. Also see Figure IA1 in the Supplementary Material for a graphical depiction of the relationship between mortgage approval rates and borrower–lender distance at small versus large banks.

The results in Table 6 indicate that small banks utilize more soft information in their lending decisions than large banks, and that this allows them to ration credit less. The results also show that soft information is most important when lending to low-income households (an intuitive result, since these households typically have less developed credit histories, which reduces the quantity and quality of hard information). Further tests in Table IA8 in the Supplementary Material show that even within lender-by-neighborhood pairs which hold location-specific information constant, small banks are more likely to lend to low-income applicants, providing evidence that the differences are rooted in their ability to assess individual borrowers rather than neighborhood factors.

The analyses here focus on the 2010 to 2015 period to preserve consistency with the results in Section IV, and to maintain the ability to use credit bureau data for control variables and to track loan performance in the next section (V.B). However, two exercises in the Supplementary Material use HMDA data to show that the differences in small versus large banks' lending are not unique to this time period. First, Figure IA2 in the Supplementary Material shows that approval rates on conventional first-lien home purchase mortgages have been higher at small banks since the start of the HMDA data in the 1990s. Second, Table IA9 in the Supplementary Material replicates the results from Table 6 using two earlier time periods. The results show that the tendency of small banks to approve more loans, and especially local loans, was even stronger during the pre-crisis period of 1995 to 2006. The results during the financial crisis period (2007–2009) are slightly weaker, but still exhibit similar patterns. Overall, these results suggest that the differences between large and small banks are driven by a fundamental component of their lending process, rather than factors specific to any particular time period.

B. Bank Size and Mortgage Loan Performance

Next, I examine loan performance to shed light on whether the higher approval rates at small banks reflect a comparative advantage lending based on soft information, or a tendency to make more "bad loans." I study loan performance using a data set that matches originated mortgages reported in the HMDA data, to loan performance information from the borrower's credit bureau file. I uniquely link loans from the two data sources based on 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. Section B of the Supplementary Material outlines the linking process in detail and provides match statistics (see Table IB1 and Figure IB1 in the Supplementary Material). I focus on conventional home purchase mortgages originated by banks from 2010 to 2013 (the tests require 2 years of post-origination performance data). I also require the property to be located in an MSA and the loan amount to be below the "jumbo" threshold. Table IB2 in the Supplementary Material summarizes this final sample of 30,954 mortgages.

The loan performance tests regress an indicator for the mortgage becoming at least 60 days delinquent during the year of origination or the following two calendar years, on an indicator for the originating bank being small. These regressions control for borrower, loan, and bank characteristics, as well as county and

TABLE 7

Mortgage Delinquencies at Small Versus Large Banks

Table 7 presents regressions of an indicator for a mortgage becoming at least 60 days delinquent in the 2 years following origination on an indicator for the loan being originated by a small bank (assets less than 1 billion in 2010 dollars). The tests control for borrower, loan, and bank characteristics, as well as county and origination month fixed effects. The mortgages are from a matched data set with information from both the credit bureau and HMDA data. Loans from these two data sources are uniquely matched based on loan characteristics (see Section B of the Supplementary Material for a detailed description of the matching process). The sample consists of conventional home purchase mortgages originated by banks from 2010 to 2013, where the property is in a metropolitan statistical area, and the loan amount is below the Government Sponsored Entities' securitization limits (excludes "jumbo" loans). Column 1 presents the results for the full sample, column 2 presents results for the subsample of low-income borrowers (below the U.S. median household income), and column 3 presents the results for high-income borrowers. Coefficients are reported in percentage point units, and the standard errors are clustered by county-year. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Full Sample	Low Income	High Income
	1	2	3
SMALL_BANK	−0.0581	−0.3392	0.0417
	(0.1986)	(0.4081)	(0.2364)
Borrower and Loan Characteristics			
CREDIT_SCORE_10-POINT_BIN_INDICATORS	Yes	Yes	Yes
NUMBER_OF_CREDIT_LINES_INDICATORS	Yes	Yes	Yes
AGE_INDICATORS	Yes	Yes	Yes
INCOME_CENTILE_INDICATORS	Yes	Yes	Yes
LOAN_AMOUNT_CENTILE_INDICATORS	Yes	Yes	Yes
LOAN_TO_INCOME_RATIO	0.0282	0.0820	0.3472
	(0.0659)	(0.0626)	(0.4581)
log(TOTAL_DEBT*t*-1)	0.0578	0.0192	0.1015
	(0.0564)	(0.1152)	(0.0626)
log(PAST_DUE_DEBT*t*-1)	0.3878***	0.3270	0.4173***
	(0.1241)	(0.2195)	(0.1517)
JOINT_MORTGAGE	−0.4608**	−0.5952	−0.3657
	(0.2308)	(0.4845)	(0.2796)
BLACK	−0.6016	−0.5464	−0.7437
	(0.7315)	(1.4633)	(0.8723)
HISPANIC	1.0327*	2.0079*	0.5201
	(0.5431)	(1.1575)	(0.5944)
Bank Characteristics			
CAPITAL_RATIO	0.1457	0.2032	0.1498
	(0.1178)	(0.2278)	(0.1393)
REAL_ESTATE_LOANS_RATIO	0.0821	−0.2183	0.2699*
	(0.1315)	(0.3077)	(0.1407)
PROFITABILITY	−0.0611	0.1127	−0.0810
	(0.1261)	(0.2685)	(0.1285)
County FE	Yes	Yes	Yes
Origination month FE	Yes	Yes	Yes
R^2	0.092	0.207	0.096
No. of obs.	30,874	7,755	22,961

origination month fixed effects. The results in Table 7 show that differences in delinquency rates at small versus large banks are statistically insignificant, and relatively small compared to the mean delinquency rate of 3.5%. These results show that small banks are able to approve a larger share of applications, without making bad loans, consistent with an information advantage (particularly in the low-income sample, where the estimated effect of SMALL_BANK on delinquency is negative).

C. Additional Evidence from Loan Sales and Rates

I next exploit the fact that the HMDA data include information on whether originated loans are sold on the secondary market. I use these data to test whether patterns in loan sales are consistent with small banks using more soft information in their lending process. A simple observation here is powerful: small banks are almost twice as likely as large banks to hold onto the loans they originate. For the sample of loans I focus on (conventional purchase mortgages for amounts below GSE securitization limits, where the property is in an MSA and within 20 miles of the bank's nearest branch) large banks retain 32% of their originations, whereas small banks retain 52%. This large difference suggests that small banks are more willing to originate difficult-to-sell mortgages (e.g., those with credit scores below GSE securitization limits) and/or that small banks' information advantage makes holding mortgages more profitable.

Additional tests in Table IA10 in the Supplementary Material examine which loans are sold. The results show that conditional on the rich set of controls, small banks are still 11.7 percentage points less likely to sell a loan. This finding suggests that roughly half of the 20-percentage-point difference between small and large banks (retaining 52% vs. 32%) can be explained by differences in loan characteristics, whereas the remainder is likely due to differences in the profitability of holding loans (even those with similar *hard information*). The cross-sectional patterns in loan sales also suggest an information-based explanation: banks are more likely to sell mortgages as the distance to the borrower increases, especially at small banks and for loans to low-income borrowers.

Lastly, I examine patterns in mortgage interest rates. The HMDA data are limited on this dimension: lenders only report the "rate spread" (the difference between the APR on the mortgage and the average rate on prime mortgages at the time) if the spread is above 1.5 percentage points.[14] I follow Bayer, Ferreira, and Ross (2018) and construct an indicator for the roughly 5% of HMDA mortgages that have a rate spread above the reporting threshold, and use this as the dependent variable. The results, also reported in Table IA10 in the Supplementary Material, show that rate spreads above the threshold are slightly more likely at small banks, and are more likely as borrower–lender distance increases (especially at small banks and for low-income borrowers). Overall, these patterns are consistent with small banks incorporating more information into their decisions, which allows them to price risk and extend credit to a larger set of borrowers.

D. Discussion of the Mechanism

1. Soft Information

The finding that small banks i) approve more mortgage applications, ii) hold onto a larger share of their originations, and iii) have similar or lower default rates on their loans suggests an information advantage in mortgage lending. The further finding that small banks' additional propensity to approve applications and retain ownership of loans is concentrated among local borrowers, and among low-income borrowers, provides evidence that their advantage is rooted in soft information. Yet, it is important to consider alternative factors such as potential differences in small versus large banks' regulatory concerns, risk management, agency problems, or underwriting flexibility. The remainder of this section discusses whether these alternatives could explain the lending patterns at small versus large banks.

[14]For details, see https://www.ffiec.gov/ratespread/newcalchelp.aspx.

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

2. Alternatives

Banks' regulatory oversight and risk management concerns have increased since the 2008 financial crisis, and it is important to consider whether heterogeneity in these concerns could explain the empirical results. Two aspects of the results run contrary to such an explanation. First, the additional approvals at small banks are concentrated among extremely local loans, which strongly supports an information-based explanation, rather than an explanation based on regulatory or risk management concerns (which would not depend on borrower–lender distance). And second, the fact that the results hold in earlier time periods shows that they are driven by fundamental differences in large versus small banks' lending process, rather than regulatory factors during a particular time period. It is also worth noting that larger banks face stronger incentives from the Community Reinvestment Act to provide credit to low-income communities (e.g., Agarwal, Benmelech, Bergman, and Seru (2012)), and yet the results throughout this paper are net of these effects.

Agency problems at small banks could explain higher approval rates for local applicants, if, for example, loan officers approved negative net present value loans to local friends or acquaintances. However, the finding that small banks approve more applications but do not experience higher defaults contradicts this explanation.

If small banks place greater authority in the hands of loan officers and branch managers, it could allow them to use more flexible underwriting criteria that are tailored to the situation, rather than rigid policies across their entire branch network. This approach could create an advantage for small banks in handling difficult to evaluate applications, such as those that do not conform to the Government Sponsored Entities' criteria, even if the evaluation is based primarily on hard information. The lines between this mechanism and using soft information are likely blurred, and both draw on the intuition from Stein (2002). However, while underwriting flexibility could contribute to small banks' higher overall approval rates, it should not directly depend on distance, and thus it still cannot explain the finding that the wedge in low-income approvals is concentrated among local loans.

In sum, the findings based on credit bureau data in Section IV, and HMDA data in Section V, show that low-income households face reduced access to credit when local banks are large, and that large banks' comparative disadvantage using soft information is a primary driver of this result.

VI. Large Banks and Intergenerational Mobility

I next examine whether the credit constraints documented above have long-run effects on children from low-income households, and hence implications for economic inequality. The seminal theory of Becker and Tomes (1979), (1986) predicts that credit constraints will limit low-income parents' investment in their children's human capital, which reduces children's chances of moving up in the income distribution, and leads to less intergenerational mobility. In practice, access to credit may facilitate a myriad of parental investments. A leading example is purchasing a home, which has been shown to improve children's living environment and educational attainment (e.g., Green and White (1997), Boehm and

Schlottmann (1999), Aaronson (2000), and Haurin, Parcel, and Haurin (2002)). Other investments in children such as private schooling, after school clubs, or postsecondary training/education can also be aided by households' access to external finance, either directly or by allowing households to smooth consumption and investment over time. Therefore, in this section, I build on the previous results and test whether having large local banks (and tighter credit constraints for low-income households) leads to lower intergenerational mobility levels.

A. Intergenerational Mobility Statistics

Studies of the determinants of intergenerational mobility have historically faced major data limitations.[15] However, Chetty et al. (2014) obtained access to administrative IRS income tax records and were allowed to compute and disseminate the first disaggregated (county-level) statistics on intergenerational mobility in the United States. I use these new mobility statistics, combined with county-level measures of LARGE_BANK_MARKET_SHARE, to evaluate the effect of having large local banks on intergenerational mobility.

The statistics from Chetty et al. (2014) describe mobility levels within a county's population in two forms. First, the authors provide the slope coefficient from a regression of children's percentile rank in the national income distribution on their parent's percentile rank. A steeper PARENT_CHILD_INCOME_SLOPE indicates lower intergenerational mobility levels in the county. Second, the authors provide quintile transition matrices describing which quintile in the national income distribution children end up in, based on the quintile their parents were in. From these data, I compute the probability that children with parents in the bottom 40% of the national income distribution transition out of the bottom 40% in terms of their incomes as adults. In later tests evaluating human capital formation, I also use slope estimates provided by Chetty et al. (2014) describing the relationship between children's college attendance and their parent's rank in the income distribution. This intermediate outcome directly measures the sensitivity of human capital formation to parental income. These county-level mobility statistics are available for one cross section based on children born between 1980 and 1982. Parental income is measured when the child is 15–19 years old, and the child's income is measured at age 26.

I collect county-level covariates from various data sources describing the county's characteristics in the year 2000. To capture the type of banks their parents had access to as the children grew up, I measure LARGE_BANK_MARKET_SHARE as of 1995 when children in the Chetty et al. (2014) data were approximately 14 years old.[16] Table 8 summarizes the county-level data set and shows that the average

[15]Data sets linking parents and children over a long enough time period to evaluate children's earnings in adulthood are scarce, and are usually based on household surveys. These data sets have small sample sizes and limited information on credit access or other factors, leading most papers to focus on accurately measuring mobility at the national level, rather than on identifying its determinants. Notable exceptions include work on the role of returns to higher education (Blanden (2009)), and on government expenditures on public schools (Mayer and Lopoo (2008)).

[16]Comprehensive data on bank branches are not available from the FDIC prior to 1994. I find similar results if I measure LARGE_BANK_MARKET_SHARE based on any year from 1994 to 2000, which covers the time period these children were in grades 6–12.

TABLE 8

Summary Statistics for Intergenerational Mobility and County Characteristics

Table 8 presents summary statistics describing intergenerational mobility levels and other characteristics for U.S. counties. The intergenerational mobility statistics and measures of income inequality are computed from IRS tax returns and published by Chetty et al. (2014). The mobility statistics use children born from 1980 to 1982 and are computed based on their income at age 26 (i.e., 2006–2008) and their parents' income when the children were 15–19 years old. The remaining county characteristics describe counties as of the year 2000, except for the information on local banks which is based on data from 1995 when the children in the Chetty et al. (2014) data were approximately 14 years old. The banking variables LARGE_BANK_MARKET_SHARE and HHI_OF_BANK_BRANCHES are based on which banks own branches in the county.

	Mean	Std. Dev.	P10	P50	P90	*N*
Intergenerational Mobility						
PARENT_CHILD_INCOME_SLOPE	0.264	0.084	0.155	0.262	0.374	2,873
TRANSITION_OUT_OF_BOTTOM_40%	0.515	0.111	0.373	0.512	0.663	2,876
PARENT_INCOME_CHILD_COLLEGE_ATTENDANCE_SLOPE	0.682	0.125	0.509	0.700	0.823	3,012
Local Banks						
LARGE_BANK_MARKET_SHARE	0.388	0.314	0.000	0.364	0.829	3,114
HHI_OF_BANK_BRANCHES	0.289	0.218	0.105	0.222	0.520	3,114
Race and Segregation						
BLACK_POPULATION_SHARE	0.086	0.141	0.001	0.017	0.306	3,138
RACIAL_SEGREGATION	0.075	0.080	0.004	0.047	0.188	3,138
SEGREGATION_OF_POVERTY	0.024	0.027	0.000	0.013	0.066	3,138
COMMUTE_LESS_THAN_15MIN	0.406	0.138	0.239	0.387	0.610	3,138
Income and Inequality						
PER_CAPITA_INCOME	32,836	6709	25,181	32,244	40,436	3,138
GINI_COEFFICIENT	0.377	0.085	0.274	0.369	0.488	3,137
TOP_1_PERCENT_INCOME_SHARE	0.094	0.044	0.050	0.083	0.150	3,036
Family Characteristics						
SINGLE_MOTHER_HOUSEHOLDS	0.194	0.066	0.124	0.182	0.278	3,138
FRACTION_OF_ADULTS_DIVORCED	0.095	0.019	0.070	0.096	0.119	3,138
FRACTION_OF_ADULTS_MARRIED	0.586	0.057	0.511	0.597	0.647	3,138
K-12 Education						
K12_STUDENT_TEACHER_RATIO	16.378	2.614	13.109	16.367	19.740	2,870
K12_TEST_SCORES_(INCOME–ADJUSTED)	−0.006	8.942	−11.761	0.761	10.436	3,089
Social Capital						
SOCIAL_CAPITAL_INDEX	−0.004	1.307	−1.645	−0.090	1.756	3,109
RELIGIOUS_POPULATION_SHARE	0.530	0.181	0.310	0.511	0.779	3,136
VIOLENT_CRIMES_PER_CAPITA	0.001	0.001	0.000	0.001	0.003	2,961
Additional Covariates						
log(POPULATION_DENSITY)	3.732	1.635	1.524	3.745	5.779	3,137
PER_CAPITA_INCOME_GROWTH_(1980–2005)	2.530	0.605	1.888	2.453	3.222	3,126

PARENT_CHILD_INCOME_SLOPE is 0.26, the probability of transitioning out of the bottom 40% of the income distribution is 51.5%, and on average, every percentile increase in parent's income rank increases the probability that their child attends college by 0.68 percentage points.

B. The Effect of Large Bank Market Share on Intergenerational Mobility

If better access to external finance with small banks helps low-income households invest in their children (e.g., by improving their environment via homeownership), then we might expect higher LARGE_BANK_MARKET_SHARE to reduce intergenerational mobility levels. However, having larger local banks may also increase lending to businesses (Demsetz and Strahan (1997)), or financial integration and economic growth (Jayaratne and Strahan (1996)). If this in turn creates jobs that facilitate mobility, then large banks' presence could instead lead to higher mobility levels. Ultimately, this is an empirical question.

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

To estimate the effect of having large local banks on intergenerational mobility, I regress county-level measures of mobility on LARGE_BANK_MARKET_SHARE, a set of 15 correlates of mobility outlined in Chetty et al. (2014), and additional controls. The 15 correlates outlined in Chetty et al. (2014) belong to five broad categories: race and segregation, income and inequality, family characteristics, kindergarten–12th grade education, and social capital. To further reduce concerns about omitted variables, I also control for population density and the growth in per capita income in the county over the lifetime of the children in the Chetty et al. (2014) data (i.e., from 1980 to 2005).

Despite this robust set of controls, concerns may remain that OLS regressions of intergenerational mobility on LARGE_BANK_MARKET_SHARE could be biased due to an omitted variables problem arising from large banks choosing where to build/buy branches. Although it is unclear in which direction we would expect any such bias to work, I use an instrumental variables approach to mitigate potential concerns. My approach exploits the staggered relaxation of state laws prohibiting interstate bank mergers from 1978 to 1997. Before the IBBEA took full effect in 1997, there was essentially no bank branching across state lines, so bank mergers were the primary way for banks to expand across state lines. Maine was the first state to open its borders to out-of-state bank entry in 1978, and by 1993 every state except Hawaii had followed suit.

In the instrumental variables approach, I use the number of years since a state opened its borders to interstate bank mergers as an instrument for LARGE_BANK_MARKET_SHARE as of 1995. The first-stage regressions (reported in Table IA11 in the Supplementary Material) show that counties in states that opened their borders to out-of-state bank entry earlier had significantly higher LARGE_BANK_MARKET_SHARE in 1995, compared to counties in states that deregulated later. This instrumental variables approach is similar in spirit to the approach employed in Berger et al. (2005), where the authors use the fraction of the prior 10 years that a state has been deregulated to instrument for local bank size.[17] The identifying assumption the approach makes is that, conditional on the county-level controls, the timing of a state's interstate banking deregulation during the 1978–1997 period influences mobility for children turning age 26 in 2007, only through its effect on the size of local banks during their childhood. This approach is supported by a long literature in finance, starting with the seminal work of Jayaratne and Strahan (1996), which treats the timing of banking deregulation as plausibly exogenous to local economic conditions (see Kroszner and Strahan (2014) for a review).

Table 9 presents the results of the OLS and instrumental variables regressions of the two measures of intergenerational mobility on LARGE_BANK_MARKET_SHARE and the controls. Columns 1–3 examine the primary measure of mobility, the log(PARENT_CHILD_INCOME_SLOPE), and columns 4–6 evaluate the second measure, TRANSITION_OUT_OF_BOTTOM_40%. Column 1 presents the baseline OLS results, which show that a standard deviation increase in LARGE_BANK_MARKET_SHARE in a county leads to a 1.98% increase in the

[17] The authors' sample is set earlier than this paper's, so the authors use the earlier wave of intrastate rather than interstate deregulation events.

TABLE 9
Large Banks and Intergenerational Economic Mobility

Table 9 presents regressions of two county-level measures of intergenerational mobility (from Chetty et al. (2014)) on the share of large bank branches in a county, and controls. Columns 1–3 present tests using the primary measure of mobility, log(PARENT_CHILD_INCOME_SLOPE), which is based on the slope coefficient from a rank–rank regression of child income centile on parent income centile within a county. The second measure, TRANSITION_OUT_OF_BOTTOM_40%, is the probability that a child with parents in the bottom 40% of the income distribution moves out of this bottom 40% as an adult. The sample is the cross section of U.S. counties for which data on all the covariates are available. Column 1 presents the OLS results, and column 2 adds state fixed effects. Column 3 presents the IV/2SLS results using YEARS_SINCE_INTERSTATE_DEREGULATION to instrument for LARGE_BANK_MARKET_SHARE (see Table IA11 in the Supplementary Material for the first-stage regressions). Columns 4–6 present similar results for the second measure of mobility. All explanatory variables are standardized to have a mean of 0 and a standard deviation of 1, and the standard errors are robust. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	log(PARENT_CHILD_INCOME_SLOPE)			TRANSITION_OUT_OF_BOTTOM_40%		
	OLS		IV	OLS		IV
	1	2	3	4	5	6
LARGE_BANK_MARKET_SHARE	0.0198***	0.0199***	0.138***	−0.00709***	−0.00113	−0.0437***
	(0.00666)	(0.00763)	(0.0363)	(0.00149)	(0.00159)	(0.00839)
Race and Segregation						
BLACK_POPULATION_SHARE	0.0898***	−0.00410	0.102***	−0.0231***	0.00311	−0.0270***
	(0.0136)	(0.0139)	(0.0153)	(0.00310)	(0.00290)	(0.00356)
RACIAL_SEGREGATION	0.0297***	0.0254***	0.0158*	−0.00356*	−0.00441***	0.000724
	(0.00784)	(0.00693)	(0.00951)	(0.00189)	(0.00145)	(0.00240)
SEGREGATION_OF_POVERTY	−0.0344***	0.0105	−0.0498***	−0.000906	−0.00824***	0.00387
	(0.00987)	(0.00903)	(0.0112)	(0.00201)	(0.00189)	(0.00247)
COMMUTE_LESS_THAN_15MIN	−0.0735***	−0.0411***	−0.0878***	0.0100***	0.00537***	0.0145***
	(0.0107)	(0.00946)	(0.0117)	(0.00222)	(0.00197)	(0.00261)
Income and Inequality						
PER_CAPITA_INCOME	−0.107***	−0.0715***	−0.133***	0.00631***	0.00159	0.0144***
	(0.0107)	(0.0102)	(0.0135)	(0.00213)	(0.00213)	(0.00301)
GINI_COEFFICIENT	0.0314**	0.00916	0.0287**	−0.0262***	−0.0159***	−0.0253***
	(0.0130)	(0.0125)	(0.0134)	(0.00267)	(0.00261)	(0.00300)
TOP_1_PERCENT_INCOME_SHARE	−0.0331***	−0.0194**	−0.0314***	0.0153***	0.00959***	0.0147***
	(0.0105)	(0.00948)	(0.0107)	(0.00223)	(0.00198)	(0.00251)
Family Characteristics						
SINGLE_MOTHER_HOUSEHOLDS	0.144***	0.130***	0.140***	−0.0437***	−0.0459***	−0.0425***
	(0.0191)	(0.0157)	(0.0205)	(0.00420)	(0.00328)	(0.00462)
FRACTION_OF_ADULTS_DIVORCED	0.0617***	0.0525***	0.0652***	−0.0174***	−0.00921***	−0.0185***
	(0.00917)	(0.00834)	(0.00954)	(0.00190)	(0.00174)	(0.00213)
FRACTION_OF_ADULTS_MARRIED	0.108***	0.0594***	0.115***	−0.0107***	−0.00184	−0.0128***
	(0.0126)	(0.0104)	(0.0135)	(0.00264)	(0.00218)	(0.00299)
K-12 Education						
K12_STUDENT_TEACHER_RATIO	−0.0373***	−0.00300	−0.0639***	−0.00210	−0.00236	0.00616**
	(0.00773)	(0.0104)	(0.0110)	(0.00154)	(0.00218)	(0.00246)
K12_TEST_SCORES_(INCOME–ADJUSTED)	0.00410	−0.0369***	0.00125	−0.00194	0.00482***	−0.00105
	(0.00884)	(0.00852)	(0.00936)	(0.00191)	(0.00178)	(0.00211)
Social Capital						
SOCIAL_CAPITAL_INDEX	0.0230**	0.0175*	0.0361***	0.0143***	0.00597***	0.0102***
	(0.00945)	(0.0104)	(0.0106)	(0.00189)	(0.00216)	(0.00232)
RELIGIOUS_POPULATION_SHARE	0.0236***	0.0258***	0.0399***	0.00873***	0.0102***	0.00369*
	(0.00805)	(0.00799)	(0.00939)	(0.00165)	(0.00167)	(0.00215)
VIOLENT_CRIMES_PER_CAPITA	0.000347	0.0158**	−0.00328	0.00375**	−0.000766	0.00487***
	(0.00642)	(0.00743)	(0.00706)	(0.00154)	(0.00155)	(0.00169)
Additional Controls						
HHI_OF_BANK_BRANCHES	0.0166	−0.0679*	0.0147	−0.0437***	−0.0117	−0.0429***
	(0.0499)	(0.0396)	(0.0527)	(0.00930)	(0.00827)	(0.0110)
log(POPULATION_DENSITY)	0.117***	−0.0145	0.1000***	−0.0240***	−0.00938***	−0.0188***
	(0.0145)	(0.0146)	(0.0161)	(0.00284)	(0.00306)	(0.00347)
PER_CAPITA_INCOME_GROWTH_(1980–2005)	−0.00284	−0.0133*	0.00334	−0.00477***	−0.000677	−0.00670***
	(0.00908)	(0.00712)	(0.00968)	(0.00170)	(0.00149)	(0.00201)
State fixed effects	No	Yes	N/A	No	Yes	N/A
R^2	0.512	0.602	–	0.763	0.814	–
No. of obs.	2,417	2,417	2,417	2,420	2,420	2,420
First-stage *F*-stat	–	–	102.1	–	–	102.5

magnitude of the PARENT_CHILD_INCOME_SLOPE, which represents a reduction in intergenerational mobility.

The tests in columns 2 and 3 of Table 9 build on this baseline result by mitigating omitted variable concerns in two different, and complementary ways. The test in column 2 adds state fixed effects, and finds that the effect of LARGE_BANK_MARKET_SHARE is virtually unchanged. This test shows that differences in various state policies outside of banking (e.g., tuition at state universities) are not acting as omitted variables and driving the relationship between the size of local banks and mobility. Then, the instrumental variables test in column 3 exploits across-state differences in the timing of interstate banking deregulation, and finds an even stronger relationship between LARGE_BANK_MARKET_SHARE and intergenerational mobility (a standard deviation increase in LARGE_BANK_MARKET_SHARE causes a 13.8% increase in the magnitude of the PARENT_CHILD_INCOME_SLOPE).[18]

This instrumental variables test uses only across-state variation due to the plausibly exogenous timing of deregulation, and therefore mitigates omitted variables that may operate within states, such as large banks choosing to build/buy branches in areas with vibrant economies and high mobility levels, which could mask any negative influence on mobility. Indeed, the fact that the instrumental variables results are stronger suggests that an omitted variable bias may be working *against* the baseline OLS results. Combined, the fixed effects test and the instrumental variables test can rule out omitted variable concerns relating to either across-state policies, or within-state differences in local economies, providing strong evidence that having large local banks reduces mobility levels.

The tests in columns 4–6 of Table 9 use the second measure of mobility, TRANSITION_OUT_OF_BOTTOM_40%, and find a similar negative effect of LARGE_BANK_MARKET_SHARE on mobility levels. Specifically, the instrumental variables estimate in column 6 shows that a standard deviation increase in LARGE_BANK_MARKET_SHARE leads to a 4.4-percentage-point decrease in the probability that a child born to parents in the bottom 40% of the income distribution transitions out of this bottom 40% in adulthood (compared to a mean of 51.5%). Overall, the results suggest that a standard deviation increase in LARGE_BANK_MARKET_SHARE leads to a reduction in intergenerational mobility levels of between 8% $\left(\frac{4.4}{51.5}\right)$ and 14%.

C. Evidence on the Homeownership and Human Capital Channels

In this final subsection, I examine the economic mechanisms underpinning the negative effect of LARGE_BANK_MARKET_SHARE on intergenerational

[18]In related work, Beck, Levine, and Levkov (2010) show that *intrastate* banking deregulation, which allowed within-state consolidation and the formation of mid-sized community banks, led to reduced income inequality. In contrast, I examine the effect of national banks entering local markets following *interstate* deregulation on the turnover/mobility within the income distribution, rather than on the shape of the income distribution itself. Based on the different changes in banking regulation and outcomes that we study, it is clear that my results are not in direct conflict with Beck et al. (2010). However, income inequality and mobility do tend to be negatively correlated, suggesting that we are identifying separate economic mechanisms.

mobility. First, I explore the cross-sectional heterogeneity in the effect, where I find evidence consistent with homeownership and other private investment in children's human capital being important channels. Then, I provide additional evidence for this human capital mechanism based on patterns in college attendance.

Given that prior tests show that large banks reduce low-income households' access to mortgage credit, and studies show that homeownership improves children's human capital formation, it is natural to examine homeownership as a potential channel through which banks affect intergenerational mobility. Indeed, purchasing a home often improves children's living environment, access to quality schools, or peer influences, and can be thought of as an investment in children's human capital.[19] However, homes with high prices are out of reach for low-income households, making the local supply of affordable homes critical for this channel (in fact, Table IA12 in the Supplementary Material shows that having small banks and hence additional credit only increases homeownership rates in counties that are above the median in affordable homes per capita).[20] Therefore, I exploit geographic variation in the supply of affordable homes to assess the importance of the homeownership channel.

Panel A of Table 10 examines the cross-sectional heterogeneity in the effect of LARGE_BANK_MARKET_SHARE on log(PARENT_CHILD_INCOME_SLOPE), by repeating the instrumental variables test on economically important subsamples. Columns 2 and 3 split the sample based on whether the county is above/below the median in terms of affordable homes per capita. The results show that the effect of LARGE_BANK_MARKET_ SHARE is over 3 times larger in areas with affordable homes, where banks and credit constraints have the largest impact on homeownership. This finding is consistent with homeownership being an important channel through which banks affect intergenerational mobility.

The second sample split examines the role of private investment in children. The supporting intuition is that government and household financial investment in children are to some extent substitutes (Solon (2004), Davies, Zhang, and Zeng (2005)). For example, high-quality public schools could alleviate some households' need to pay for private primary and secondary education. To capture areas with more/less government investment in children, I split the sample based on the student–teacher ratio in local K–12 public schools. The results in columns 4 and 5 of Table 10 show that LARGE_BANK_MARKET_SHARE has a much larger effect on mobility levels in areas with less government investment (higher student–teacher ratios), where more of the financial burden tends to be placed on households. This pattern is consistent with credit constraints and reduced private investment in children being a channel through which large banks affect mobility.

The final set of tests provide an additional, fairly direct, piece of evidence that a mechanism involving children's human capital formation is indeed at work. The

[19]For evidence that homeownership improves children's home environment and educational attainment, see Green and White (1997), Boehm and Schlottmann (1999), Aaronson (2000), and Haurin et al. (2002).

[20]I define affordable homes per capita as the number of owner-occupied homes valued at or below \$100,000 divided by the county population.

TABLE 10

Evidence on the Homeownership and Human Capital Channels

Table 10 presents evidence on the human capital channel by examining cross-sectional heterogeneity in the effect of large banks on mobility, and by examining patterns in college attendance. Panel A presents IV/2SLS estimates of the effect of LARGE_BANK_MARKET_SHARE in a county on intergenerational mobility, as measured by log(PARENT_CHILD_INCOME_SLOPE), for important subsamples. 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 (see Table IA11 in the Supplementary Material for the first-stage regressions). 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, for reference. Columns 2–5 show sample splits for counties above/below the median levels of affordable owner-occupied homes per capita, and student–teacher ratios in local schools. Panel B presents regressions evaluating whether children's human capital attainment is more sensitive to parental income in areas where local banks are large. The dependent variable is the log(PARENT_INCOME_CHILD_COLLEGE_ATTENDANCE_SLOPE), which is based on the coefficient from a regression of an indicator for children's college attendance on their parent's rank in the national income distribution, computed for residents of a given county by Chetty et al. (2014). Column 1 presents the OLS results, column 2 presents state fixed effects, and column 3 presents the IV/2SLS results. Explanatory variables are standardized to have a mean of 0 and a standard deviation of 1, and the standard errors are robust. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

Panel A. Large Banks and the Log(PARENT_CHILD_INCOME_SLOPE) – IV Results for Subsamples

	Full Sample	Affordable Homes		Student–Teacher Ratio	
		High	Low	High	Low
	1	2	3	4	5
LARGE_BANK_MARKET_SHARE	0.138***	0.228***	0.0676*	0.138**	0.0720
	(0.0363)	(0.0750)	(0.0400)	(0.0609)	(0.0478)
All controls	Yes	Yes	Yes	Yes	Yes
No. of obs.	2,417	1,212	1,205	1,264	1,153
First-stage *F*-stat	102.1	32.7	62.8	34.1	59.2

Panel B. Large Banks and the Sensitivity of College Attendance to Parental Income

	OLS		IV
	1	2	3
LARGE_BANK_MARKET_SHARE	0.00386	0.0117**	0.158***
	(0.00424)	(0.00455)	(0.0251)
All controls	Yes	Yes	Yes
State fixed effects	No	Yes	N/A
R^2	0.346	0.522	–
No. of obs.	2,533	2,533	2,533
First-stage *F*-stat	–	–	119.2

tests examine patterns in the sensitivity of children's college attendance to their parents' income. A college education is a critical component of human capital in the long run, and the decision to attend college is a strong indicator for a child's human capital at age 18–19. Therefore, if large banks lead to tighter credit constraints and reduced investment in children from low-income households, we should expect to see a greater sensitivity of college attendance to parental income. Importantly, these tests can establish whether large banks' effect can be seen in children even before they enter the labor market, which would be consistent with a mechanism rooted in parental investments rather than alternative explanations such as banks affecting mobility through an influence on local companies or labor markets.

Panel B of Table 10 examines the relationship between the size of local banks and the natural logarithm of the PARENT_INCOME_CHILD_COLLEGE_ATTENDANCE_SLOPE, which is the slope coefficient from a regression of an indicator for children's college attendance at age 19 on their parents' rank in the national income distribution, computed for the residents of a given county by Chetty et al. (2014). The baseline OLS results in column 1 show a positive but statistically insignificant effect of LARGE_BANK_MARKET_SHARE.

https://doi.org/10.1017/S0022109023001114 Published online by Cambridge University Press

However, once omitted variable concerns are addressed more directly, either by including state fixed effects (column 2), or by employing the instrumental variables approach based on the timing of deregulation (column 3), a clear finding emerges: having large local banks increases the sensitivity of college attendance to parental income.[21]

Overall, the findings in Table 10 suggest that having large local banks (and tighter credit constraints) reduces human capital formation in low-income households, likely by limiting homeownership and other forms of private investment in children. In line with Becker and Tomes (1979), (1986), this reduced investment in children from low-income households ultimately reduces intergenerational mobility levels. As novel micro-data become available in the future, perhaps these and other aspects of financial institutions' influence on economic mobility can be explored even further.

VII. Conclusion

This paper finds that the structure of the banking industry affects the distribution of credit across households. When local banks are large, borrowers with low incomes, subprime credit scores, and/or limited credit histories experience reduced access to credit. In contrast, borrowers of high economic status continue to receive credit. These fundamental results hold in a variety of subsamples, and when employing an instrumental variables approach that exploits regulatory differences across state borders. Further tests show that large banks utilize less soft information when lending to households, and that this information is especially important when lending to low-income households. These findings offer an explanation for why low-income households face reduced credit access when large banks dominate the local market.

The asymmetric effect of the size of local banks on high- versus low-income households leads this paper to consider broader effects on economic inequality. I test the theoretical prediction from Becker and Tomes (1979), (1986) that access to credit should facilitate low-income households' investment in their children's human capital, and foster intergenerational economic mobility. Indeed, I find that having large local banks leads to lower intergenerational mobility levels. Both the cross-sectional patterns, and tests based on children's educational attainment, suggest that this result is driven primarily by low-income households' credit constraints and human capital formation. These findings provide the first evidence of a link between the structure of the banking industry and intergenerational economic mobility. Further exploration of the determinants of mobility, including the role of credit constraints and financial institutions, is a promising avenue for future research.

[21]The instrumental variables approach ruling out omitted variables within states is likely important here, because it seems plausible that large banks might tend to build/buy branches in parts of a state that are thriving, where children from low-income families are more likely to attend college, which could mask any negative effect on low-income households. Indeed, the stronger instrumental variables results suggest that such an omitted variable may be working against the OLS results.

Supplementary Material

To view supplementary material for this article, please visit http://doi.org/10.1017/S0022109023001114.

References

Aaronson, D. "A Note on the Benefits of Homeownership." *Journal of Urban Economics*, 47 (2000), 356–369.

Agarwal, S.; B. W. Ambrose; S. Chomsisengphet; and C. Liu. "The Role of Soft Information in a Dynamic Contract Setting: Evidence from the Home Equity Credit Market." *Journal of Money, Credit and Banking*, 43 (2011), 633–655.

Agarwal, S.; E. Benmelech; N. Bergman; and A. Seru. "Did the Community Reinvestment Act Lead to Risky Lending?" NBER Working Paper No. 18609 (2012).

Agarwal, S., and R. Hauswald. "Distance and Private Information in Lending." *Review of Financial Studies*, 23 (2010), 2757–2788.

Akey, P.; C. Dobridge; R. Z. Heimer; and S. Lewellen. "Pushing Boundaries: Political Redistricting and Consumer Credit." Working Paper, University of Toronto (2018).

Akey, P.; R. Z. Heimer; and S. Lewellen. "Politicizing Consumer Credit." *Journal of Financial Economics*, 139 (2021), 627–655.

Avery, R. B.; M. F. Bilinski; B. K. Bucks; C. Chai; T. Critchfield; I. H. Keith; I. E. Mohamed; F. W. Pafenberg; S. Patrabansh; J. D. Schultz; and C. E. Wood. "A Profile of 2013 Mortgage Borrowers: Statistics from the National Survey of Mortgage Originations." Technical Report (2017)

Bayer, P.; F. Ferreira; and S. L. Ross. "What Drives Racial and Ethnic Differences in High-Cost Mortgages? The Role of High-Risk Lenders." *Review of Financial Studies*, 31 (2018), 175–205.

Beck, T.; R. Levine; and A. Levkov. "Big Bad Banks? The Winners and Losers from Bank Deregulation in the United States." *Journal of Finance*, 65 (2010), 1637–1667.

Becker, G. S., and N. Tomes. "An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility." *Journal of Political Economy*, 87 (1979), 1153–1189.

Becker, G. S., and N. Tomes. "Human Capital and the Rise and Fall of Families." *Journal of Labor Economics*, 4 (1986), S31–S39.

Berger, A. N., and L. K. Black. "Bank Size, Lending Technologies, and Small Business Finance." *Journal of Banking and Finance*, 35 (2011), 724–735.

Berger, A. N.; C. H. Bouwman; and D. Kim. "Small Bank Comparative Advantages in Alleviating Financial Constraints and Providing Liquidity Insurance over Time." *Review of Financial Studies*, 30 (2017), 3416–3454.

Berger, A. N.; N. H. Miller; M. A. Petersen; R. G. Rajan; and J. C. Stein. "Does Function Follow Organizational Form? Evidence from the Lending Practices of Large and Small Banks." *Journal of Financial Economics*, 76 (2005), 237–269.

Berger, E. A.; A. W. Butler; E. Hu; and M. Zekhnini. "Financial Integration and Credit Democratization: Linking Banking Deregulation to Economic Growth." *Journal of Financial Intermediation*, 54 (2021), Article no. 100857.

Bhutta, N., and B. J. Keys. "Interest Rates and Equity Extraction During the Housing Boom." *American Economic Review*, 106 (2016), 1742–1774.

Black, S. E., and P. J. Devereux. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, 4th edition, D. Card and O. Ashenfelter, eds. Amsterdam: Elsevier (2011), 1487–1541.

Blanden, J. "How Much Can We Learn from International Comparisons of Intergenerational Mobility?" Working Paper, Centre for the Economics of Education (2009), 1–45.

Boehm, T. P., and A. M. Schlottmann. "Does Home Ownership by Parents Have an Economic Impact on Their Children?" *Journal of Housing Economics*, 8 (1999), 217–232.

Bord, V. M. "Bank Consolidation and Financial Inclusion: The Adverse Effects of Bank Mergers on Depositors." Working Paper, Harvard University (2018).

Brown, J. R.; J. A. Cookson; and R. Z. Heimer. "Growing Up Without Finance." *Journal of Financial Economics*, 134 (2019), 591–616.

Butler, A. W.; E. J. Mayer; and J. P. Weston. "Racial Disparities in the Auto Loan Market." *Review of Financial Studies*, 36 (2023), 1–41.

Carter, D. A., and J. E. McNulty. "Deregulation, Technological Change, and the Business-Lending Performance of Large and Small Banks." *Journal of Banking and Finance*, 29 (2005), 1113–1130.

Celerier, C., and A. Matray. “Bank-Branch Supply, Financial Inclusion, and Wealth Accumulation.” *Review of Financial Studies*, 32 (2019), 4767–4809.

Cetorelli, N., and P. E. Strahan. “Finance as a Barrier to Entry: Bank Competition and Industry Structure in Local U.S. Markets.” *Journal of Finance*, 61 (2006), 437–461.

Chava, S.; A. Oettl; A. Subramanian; and K. V. Subramanian. “Banking Deregulation and Innovation.” *Journal of Financial Economics*, 109 (2013), 759–774.

Chetty, R.; N. Hendren; P. Kline; and E. Saez. “Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States.” *Quarterly Journal of Economics*, 129 (2014), 1553–1623.

Cole, R. A.; L. G. Goldberg; and L. J. White. “Cookie Cutter vs. Character: The Micro Structure of Small Business Lending by Large and Small Banks.” *Journal of Financial and Quantitative Analysis*, 39 (2004), 227–251.

Cornaggia, J.; K. J. Cornaggia; and H. Xia. “College Student Behavior and Student Loan Default.” Working Paper, Pennsylvania State University (2021).

Davies, J. B.; J. Zhang; and J. Zeng. “Intergenerational Mobility Under Private vs. Public Education.” *Scandinavian Journal of Economics*, 107 (2005), 399–417.

Demsetz, R. S., and P. E. Strahan. “Diversification, Size, and Risk at Bank Holding Companies.” *Journal of Money, Credit, and Banking*, 29 (1997), 300–313.

DeYoung, R.; D. Glennon; and P. Nigro. “Borrower-Lender Distance, Credit Scoring, and Loan Performance: Evidence from Informational-Opaque Small Business Borrowers.” *Journal of Financial Intermediation*, 17 (2008), 113–143.

Diamond, D. W. “Financial Intermediation and Delegated Monitoring.” *Review of Economic Studies*, 51 (1984), 393–414.

Gaviria, A. “Intergenerational Mobility, Sibling Inequality and Borrowing Constraints.” *Economics of Education Review*, 21 (2002), 331–340.

Green, R. K., and M. J. White. “Measuring the Benefits of Homeowning: Effects on Children.” *Journal of Urban Economics*, 41 (1997), 441–461.

Haurin, D. R.; T. L. Parcel; and R. J. Haurin. “Does Homeownership Affect Child Outcomes?” *Real Estate Economics*, 30 (2002), 635–666.

Iyer, R.; A. I. Khwaja; E. F. P. Luttmer; and K. Shue. “Screening Peers Softly: Inferring the Quality of Small Borrowers.” *Management Science*, 62 (2016), 1554–1577.

Jayaratne, J., and P. E. Strahan. “The Finance-Growth Nexus: Evidence from Bank Branch Deregulation.” *Quarterly Journal of Economics*, 111 (1996), 639–670.

Kroszner, R. S., and P. E. Strahan. “Regulation and Deregulation of the US Banking Industry: Causes, Consequences, and Implications for the Future.” In *Economic Regulation and Its Reform: What Have We Learned?* N. L. Rose, ed. Chicago, IL: University of Chicago Press, (2014), 485–543.

Lee, D., and W. van der Klaauw. “An Introduction to the FRBNY Consumer Credit Panel.” Federal Reserve Bank of New York Staff Report No. 479 (2010).

Loutskina, E., and P. E. Strahan. “Securitization and the Declining Impact of Bank Finance on Loan Supply: Evidence from Mortgage Originations.” *Journal of Finance*, 64 (2009), 861–889.

Loutskina, E., and P. E. Strahan. “Informed and Uninformed Investment in Housing: The Downside of Diversification.” *Review of Financial Studies*, 24 (2011), 1447–1480.

Mayer, S. E., and L. M. Lopoo. “Government Spending and Intergenerational Mobility.” *Journal of Public Economics*, 92 (2008), 139–158.

Mazumder, B. “Fortunate Sons: New Estimates of Intergenerational Mobility in the United States Using Social Security Earnings Data.” *Review of Economics and Statistics*, 87 (2005), 235–255.

Minton, B. A.; A. G. Taboada; and R. Williamson. “Bank Mergers, Acquirer Choice and Small Business Lending: Implications for Community Investment.” NBER Working Paper No. 29284 (2021).

Munnell, A. H.; G. M. B. Tootell; L. E. Browne; and J. McEneaney. “Mortgage Lending in Boston: Interpreting HMDA Data.” *American Economic Review*, 86 (1996), 25–53.

Petersen, M. A., and R. G. Rajan. “Does Distance Still Matter? The Information Revolution in Small Business Lending.” *Journal of Finance*, 57 (2002), 2533–2570.

Rice, T., and P. E. Strahan. “Does Credit Competition Affect Small-Firm Finance?” *Journal of Finance*, 65 (2010), 861–889.

Rupasingha, A.; S. J. Goetz; and D. Freshwater. “The Production of Social Capital in US Counties.” *Journal of Socio-Economics*, 35 (2006), 83–101.

Solon, G. “A Model of Intergenerational Mobility Variation over Time and Place.” In *Generational Income Mobility in North America and Europe*, M. Corak, ed. Cambridge: Cambridge University Press, (2004), 38–47.

Stein, J. C. “Information Production and Capital Allocation: Decentralized Versus Hierarchical Firms.” *Journal of Finance*, 57 (2002), 1891–1921.

Stiglitz, J. E., and A. Weiss. "Credit Rationing in Markets with Imperfect Information." *American Economic Review*, 71 (1981), 393–410.
Strahan, P. E., and J. P. Weston. "Small Business Lending and the Changing Structure of the Banking Industry." *Journal of Banking and Finance*, 22 (1998), 821–845.