August 7, 2023

Supplement Filing Regarding the Community Reinvestment Act (CRA) Proposed Rule Reviewing Fed Data Demonstrating That the CRA Rule Will Not Work and Redlining Will Continue

Jerome H. Powell, Federal Reserve Board Chair
Michael S. Barr, Federal Reserve Board Vice Chair Supervision
Michelle W. Bowman, Federal Reserve Board Governor
Lisa D. Cook, Federal Reserve Board Governor
Phillip N. Jefferson, Federal Reserve Board Governor
Christopher J. Waller, Federal Reserve Board Governor

Martin J. Gruenberg, Federal Deposit Insurance Corporation Chair
Travis Hill, Federal Deposit Insurance Corporation Vice Chairman
Jonathan McKernan, Federal Deposit Insurance Corporation Director
Rohit Chopra, Federal Deposit Insurance Corporation *ex officio* Director

Michael J. Hsu, Acting Comptroller of the Currency

Dear Members of the Board of Governors, Directors of the FDIC, and Acting Comptroller of the Currency:

We have undertaken extensive statistical analysis of the proposed Community Reinvestment Act (CRA) rule set forth in the May 2022 Notice of Proposed Rulemaking (NPR) and tested it against the Federal Reserve's ("Fed") historical CRA data. This analysis has demonstrated grave deficiencies in the proposed rule. These deficiencies will perpetuate and accentuate the significant failings of the current rule, which has rewarded banks with high CRA ratings even as they steadily exited retail lending while increasing their deposit bases. The proposed rule will do nothing to change these trends and conduct, which are contrary to the letter and intent of the CRA. In fact, the proposed rule will amplify indefensible deficiencies of the current rule. Most shockingly, neither rule is equipped to detect textbook cases of redlining, which is supposed to be the focus of the CRA and addressed by the rules enacting the CRA. That is a glaring failure that must be fixed before the NPR is finalized.

We presented detailed findings to explain this analysis to the Federal Deposit Insurance Corporation ("FDIC") staff. We also tried to do the same for the Fed staff, but they would only accept a summary letter. We were specifically told that filing these findings in a supplemental comment letter with the banking agencies responsible for the CRA "would not be helpful." While data and facts showing that the proposed rule in the NPR does not work may "not be helpful," it is much more important for LMI communities to finally have a Community Reinvestment Act that works for them and delivers the concrete results promised to them for many decades now. Thus far, it has worked to bestow upon banks high CRA ratings while homeownership among LMI communities has not improved in the 45 years since the CRA was first enacted (see Figure 8 below).

Redlining, de jure and de facto, is far too polite a word for the egregious harm done to LMI communities for about 100 years. Redlining is an intentional and illegal restriction on the flow of capital that creates, exacerbates, and perpetuates inequality, poverty, and racism, often intergenerationally. The result is too few loans for homes or small businesses and little if any equity or wealth building or investment, causing a downward spiral in housing values, erosion of the tax base, deterioration of community infrastructure including schools, and destruction of neighborhoods. Many lives and dreams have been and continue to be destroyed by this illegal conduct. First imposed by the government, it was perpetuated by the private sector with banks and financial companies in the lead. Now is the time to change that by finalizing a new CRA rule that will actually make a meaningful difference in the lives of the people that live in LMI communities across this country. This is the statutory mandate that the banking agencies are required to fulfill.

We are aware of no comprehensive statistical analysis of the NPR other than that which is presented herein. The NPR presented just one summary table. Among the approximately 500 comment letters submitted in response to the NPR, there was only one attempt at a quantitative assessment of the Retail Lending Test (RLT)[1], and no such analysis has appeared since to our knowledge. While it is not clear why the banking agencies did not do such analysis or why it was not part of the NPR, it is not surprising that no other commentator has done this given the thousands of hours and expertise in multiple subjects necessary to undertake this analysis, which included parsing the needlessly impenetrable NPR, creating a model, inputting massive amounts of data, doing the analysis, and creating numerous tables, graphs and charts, several of which are presented herein.

We wish we had the resources to produce this analysis sooner and to supplement our comments earlier, but we are a very small nonprofit with limited resources. We did this as quickly as we could, tried for some time to present it to your staff, and, failing to get the response the data and analysis warrants (indeed, getting the opposite), we are hereby filing this to supplement our comment letter that we filed on August 2, 2022.[2]

---

1 Laurie Goodman, Linna Zhu, Jun Zhu, Ellen Seidman, John Walsh, Janneke Ratcliffe, *Community Reinvestment Act Modernization, Comments on the May 2022 Notice of Proposed Rulemaking*, Urban Institute (August 2022), https://www.urban.org/research/publication/community-reinvestment-act-modernization This analysis compares banks in aggregate to nonbanks in each assessment area, and thus abstracts from many of the features that determine results for individual banks.

2 *See* Better Markets' comment letter (Aug. 5, 2022), available here (including *Appendix: Analysis of CRA Related Data* starting at p. 30).

In that prior letter, we emphasized the need to benchmark the new CRA performance evaluation framework laid out in the NPR. Since that time, we have replicated the RLT, applied it to the historical CRA dataset posted on the Fed website, and compared its conclusions with alternative tests of bank lending performance. This quantitative analysis has uncovered features of the proposed RLT that defeat the stated purpose of the proposed rule in the NPR, and identifies lines along which it can be improved, including the following:

1. The RLT preserves the "static" perspective of its predecessor that rendered it blind to banks' wholesale decline in LMI lending and loan-to-deposit ratios over the last decade, and its "local" perspective that is functionally blind to redlining. It also gives a more prominent role to components of the existing performance evaluation framework – Community Benchmarks and the "best-of" clause[3] – which do not operate in line with the economic justification offered for them.
2. The vast majority of banks' RLT failure rate traces to secondary components of the RLT, or those whose economic justification does not withstand scrutiny. The design and calibration of these components will likely prompt unintended consequences from banks, in their efforts to pass the RLT, to the detriment of LMI communities.
3. A simpler version of the test that excludes the Community Benchmark, the "best of" clause, and the 60% Test produces much the same aggregate failure rate outcome as the RLT but lessens potentially adverse incentives for banks. This version is restricted to comparisons of bank performance just with the Market Benchmark, with or without the addition of the Retail Lending Volume Screen (RLVS).
4. This simpler test is directly comparable with a statistical benchmark, which tests the hypothesis that a bank's LMI lending share differs from the market only by random error. The NPR's "benchmark multipliers" are versions of statistical confidence levels from this perspective. Benchmarking against the explicit statistical test reveals that, by conventional statistical standards, the benchmark multipliers are quite lenient in their assignment of failure decisions.
5. We suggest additional features and adjustments that align the new test better with CRA objectives. Consideration of potential demand for loans, which motivated reliance on Community Benchmarks, is reintroduced by scaling results for each Assessment Area by a factor that depends on aggregate LMI loans per LMI family. This incentivizes banks to move their LMI lending efforts to less-served areas. To make the test register banks' exit from home lending, the most feasible avenue is to reformulate the Retail Lending Volume Screen as a requirement for an absolute dollar amount of loans per dollar of deposits.

Each of these findings is substantiated and reviewed in this letter.

---

[3] The "best-of" clause is our term for the component of the rule that compares a banks LMI lending share with the most favorable of Community or Market Benchmark, each scaled by their respective multipliers (see Table 4).

1. The RLT preserves the "static" perspective of its predecessor that rendered it blind to banks' wholesale decline in LMI lending and loan-to-deposit ratios over the last decade, and its "local" perspective that is functionally blind to redlining.

a) Limitations of the existing Lending Test

Retail lending by banks, especially to LMI communities, is a central concern of the CRA. For example:

> The Community Reinvestment Act was enacted in 1977, against a backdrop of urban decay and a lack of investment in communities. Congress found that banks have a continuing and affirmative obligation to help meet the credit needs of their local communities, including low- and moderate-income (LMI) neighborhoods where they are chartered, consistent with the safe and sound operations of the institutions. This finding was based on preexisting chartering laws that require banks to demonstrate that their deposit taking facilities serve the convenience and needs of their communities, which include credit and deposit services.[4]

Evaluations of the CRA published fifteen-or-so years ago[5] pointed to the success of the Act in materially increasing the number of LMI loans made by banks. They also noted a qualitative change: the CRA had contributed to making LMI lending a going concern for banks, whereas before "because of racial discrimination or fear of credit weakness, many banks 'redlined' entire areas of American cities as places they would not lend".[6]

After the Great Recession, banks' LMI lending and their loan-to-deposit ratios both dropped precipitously, and only recovered slightly by 2019, according to the CRA data. Figure 1 charts the paths of banks' deposits and home mortgage originations covered by CRA. Evidently, while deposits rose steadily, loans declined overall.

---

4 https://www.federalreserve.gov/consumerscommunities/cra_history.htm. The OCC's statement begins: "The Community Reinvestment Act (CRA) was enacted in 1977 to prevent redlining…" and continues in a similar vein.

5 *The CRA: past successes and future opportunities*, (2009) Eugene A. Ludwig, James Kamihachi, and Laura Toh, in *Revisiting the CRA: Perspectives on the Future of the Community Reinvestment Act*, https://www.frbsf.org/community-development/wp-content/uploads/sites/3/cra_past_successes_future_opportunities1.pdf

6 Ludwig et al., p.84

Figure 1: CRA-eligible Bank Loans and Deposits

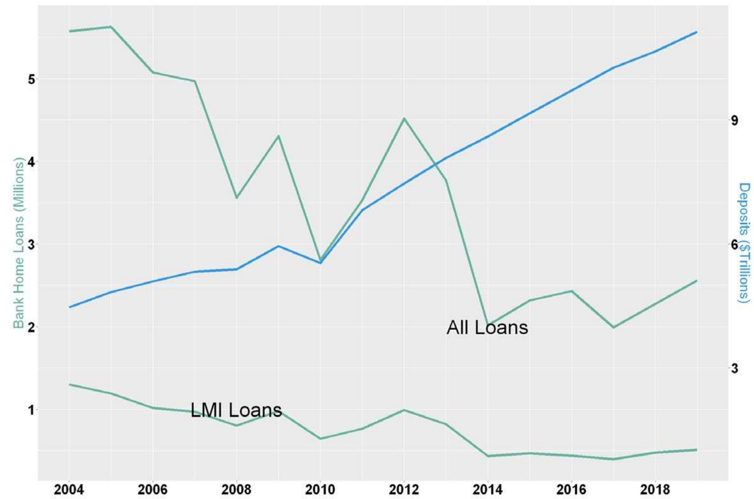
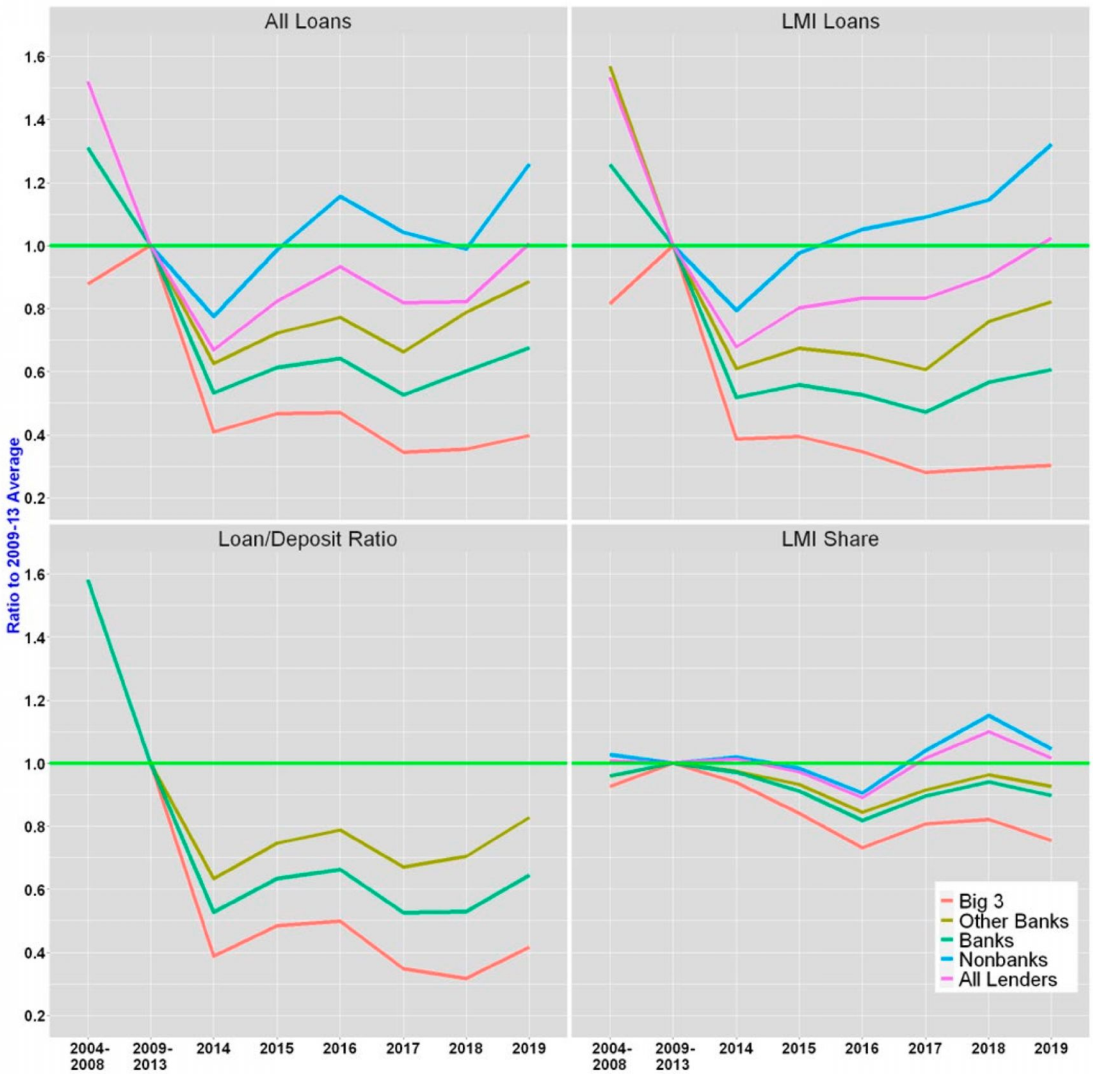


Figure 2 breaks out the lending statistics for banks and nonbanks, and further disaggregates into the Big 3[7] and other banks. To make comparisons easier, each series is shown relative to its average for the years following the Great Recession, 2009-13 (or “par”, for short). Even after this adjustment, a precipitous decline in lending across the board in 2014 is still apparent.[8]

7 JP Morgan Chase, Bank of America, and Wells Fargo.

8 Interest rates rose about 1.5% that year, although there was no slowdown in economic activity. Neil Bhutta, Steven Laufer and Daniel R. Ringo, (2017) “The Decline in Lending to Lower-Income Borrowers by the Biggest Banks”, discusses banks’ concerns about incurring triple damages on FHA loans as a possible catalyst of the decline in LMI lending.

Figure 2: Lending History Relative to 2009-13 Average ("Par")



By 2019, while bank lending had barely budged from its low, the market recovered to par, the slack taken up by nonbanks.[9] The Big 3 display the most extreme version of this pattern. Their loan volumes continued to shrink after 2014; by 2019 the Big 3's total loans had shrunk to 40% of their par value, while LMI lending was at 30% of par.[10] Because both Total and LMI loans declined by the same order of magnitude, the movements in the LMI share of loans are less marked, and all but the Big 3 are close to par by 2019. The bottom-left panel of Figure 2 charts a similar decline in banks' loans in relation to deposits.

Banks' contraction of home loans thus reversed the advances made before the Great Recession. Banks extended fewer LMI loans and lent less per dollar of deposits, both of which are

---

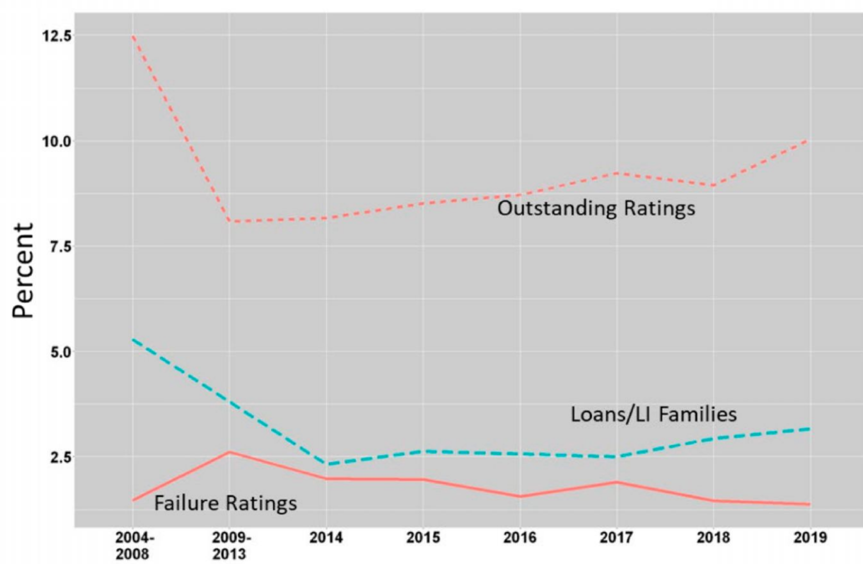
9 From par to 2019, bank lending declined from 43% to 29% of the CRA Market and also declined about one-third in aggregate.

10 Low-Income and Moderate-Income loans are combined here, for simplicity. Declines were greater in the Low-Income category than in the LMI aggregate.

inconsistent with their "affirmative obligation to help meet the credit needs of their local communities, which include credit and deposit services."[11] However, over the last two decades banks' CRA performance evaluations have registered no change, and the annual failure rate[12] has hovered between 1.6% and 3.5% by "headcount," with failing banks consistently accounting for about 10% of all banks' asset value.

Although banks' CRA footprint shrunk, it does not automatically follow that they were not helping to meet the credit demands of their communities. Perhaps they were successfully meeting the residual demand, after nonbanks had made their loans. This is a hard case to make, especially for low-income loans, whose volume shrank in absolute terms from par to 2019, while the number of low-income families grew by 8%.[13] Figure 3 illustrates the consequent path of loans per LI Family, comparing it with the low and steadily declining CRA exam failure rate, and the steadily increasing "Outstanding" rate after the Great Recession.

Figure 3: CRA Exam Ratings and LI Loans per Family
(Ratings are expressed as the share of banks rated in each year by headcount)



Why didn't banks' CRA performance evaluations worsen as their LMI lending contracted? The current Lending Test, which makes up 50% of a bank's performance evaluation, compares lending performance across lenders at a point in time, *not* over time. One component of the Lending Test, the "Distribution of Loans by Income Level of the Borrower/Geography",

---

11 Another argument that points the finger at declining demand rather than supply of loans invokes the decline in subprime loans after the Great Recession. However, while the subprime market shrank, its overlap with banks' CRA Assessment Areas was minimal. See Reid, C., Seidman, E., Willis, M., Ding, L., Silver, J. and Ratcliffe, J. (2013), "Debunking the CRA Myth—Again." Center for Community Capital, University of North Carolina at Chapel Hill.

12 In this document, a rating of Substantial Noncompliance or Needs to Improve is called a failure.

13 Calculated from https://www.ffiec.gov/craratings/craratng.zip. See also Darryl E. Getter, *The Effectiveness of the Community Reinvestment Act*, Congressional Research Service, 2015.

benchmarks a bank's LMI lending share against the market LMI lending share at the same time. ***This test is blind to a bank lowering its total and LMI lending proportionately***.

Another component, the "Lending Activity" test, benchmarks a bank's lending in relation to deposits against all other banks' contemporaneous loan/deposit ratio. Here, the ***blind spot*** is all banks lowering their home lending uniformly, even though this can involve a dramatic change in the mix of "credit and deposit services" they offer. When the major achievement identified before the Great Recession--the growth in banks' LMI lending—is reversed, it simply doesn't register, because all banks have also lowered their total lending roughly proportionately.

The Lending Test will not detect redlining by all banks in a geography, for the same reasons as given above: the test is blind to practices shared by all lenders. Each regulator's website, and almost every article on the CRA identifies eliminating redlining as one of its most pressing goals, but the CRA rule is simply blind to it.

<u>b) The RLT and its similarities with the original Lending Test:</u>

For clarity, we lay out our high-level understanding of the RLT recipe, and introduce specific names for certain components of the RLT that are not named explicitly in the NPR

- The focal statistic is the bank's share of loans made in a low-income area (LI) or a moderate-income area (MI) relative to its total loans in a geographic Assessment Area (AA). This share is compared with two benchmarks for the AA:
  - The Market Benchmark (MB), which is the corresponding loan ratio for all lenders in the AA.
  - The Community Benchmark (CB), a demographic designation of the number of potential borrowers in an income group, designated either by "geography" or "borrower."
- The "Score Test" assigns a point score to the comparisons of the bank's LI or MI share with each benchmark, using a scale of "benchmark multipliers" (see Table 4). The bank receives the higher of the two point scores, which is called the "best-of clause" in this letter.
- The "Retail Lending Volume Screen" caps the bank's point score in the AA if its loan/deposit ratio is less than 30% of the ratio of all other banks in the AA.
- The bank's aggregate score is the average of its scores across all AAs, each weighted by a 50-50 mix of its share in the bank's total loans and total deposits. This aggregate score is capped if the bank fails to exceed a threshold score in at least 60% of its AAs by headcount, which this letter calls the "60% Rule." This rule is applied to the bank's score in the AA across the aggregate of Retail Lending (45%), Community Development Financing (30%), Retail Service and Products (15%) and Community Development Services (10%) Tests.
- The bank's aggregate score is mapped into a grid of ratings for the RLT. The ratings are "Outstanding," "High Satisfactory," "Low Satisfactory," "Needs Improvement," and "Substantial Noncompliance," which are abbreviated here to OS, HS, LS, NI and SN. In this letter, a rating of SN or NI is called a failure.

The Score Test corresponds to the Lending Test's "Distribution of Loans by Income Level of the Borrower/Geography." While the CB was invoked only in specific instances in the Lending Test, it is part of the Score Test in every AA. The Retail Lending Volume Screen takes the place of the Lending Activity test. Both tests are conducted over an average of years, and so preserve the static nature of the Lending Test, and its inability to detect changes in the volume of loans that occur uniformly across banks. They also preserve the local nature of the Lending Test, in that they compare only lending activity within the same geography, and so have no way of flagging redlining by all banks in concert.

c) Problems With CBs and the "Best-of" Clause

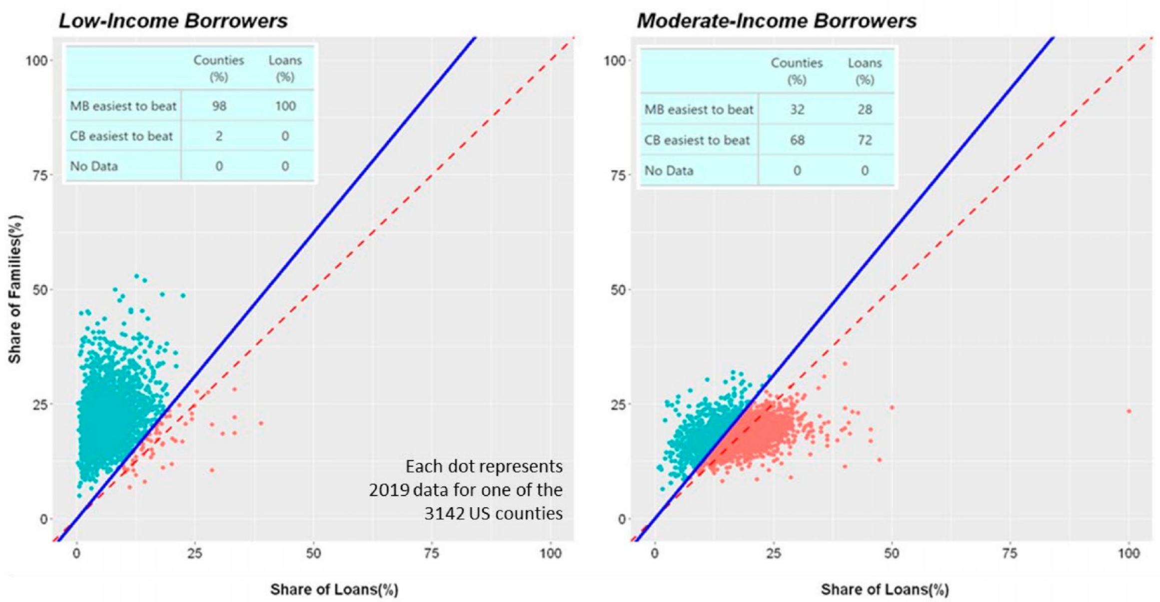
Certain features that were deployed only occasionally with discretion in the original Lending Test –CBs and the "best-of" clause-- are used in every Score Test in the RLT. Their economic rationale receives limited attention in the NPR. In several places the NPR asserts that a CB "measures" potential demand while the MB measures the amount of lending that occurred. (See, e.g., NPR at p.204). This claim about the CB is true only when there is one type of borrower across all income groups, or all income groups have the same proportionate mix of borrower types.[14] In the general case, the CB is unrelated to the unobserved demand for loans, and there is no reason to believe that the best-of clause enforces benchmarking to potential demand.

To illustrate, say each income group comprises "demand groups," each with its own propensity to demand loans. To take a life-cycle perspective, demand groups might be renters, growing families, empty-nesters, and retirees. One in ten renters could be in the loan market as a first-time buyer each year, whereas the demand propensity of empty-nesters could be that one in five look to downsize each year, and so are in the market for a purchase loan. From one geography to the next, each demand group can also make up a different proportion of an income group population. Last, lenders may also have different propensities to lend to each group, for reasons of safety and soundness, or discrimination.

14 Denote the number of families in demand group d and income group i by N_{id}, where an income group is LI or MI or those above MI, and a demand group is as laid out in the text. Each demand group has a potential demand p_{id}, which measures the probability that a family in group *id* will be in the market for a loan in the year in question. Then the families-based CB for income group i is $C_i = \sum_d N_{id} / \sum_{jd} N_{jd}$, and the corresponding measure of potential demand is $D_i = \sum_d N_{id}\, p_{id} / \sum_{jd} N_{jd}\, p_{jd}$. These two will be equal if p_{jd} is the same for all j and d, or if p_{jd} is the same for all j and $N_{id} = m_i N_d$ (i.e., each income group is composed of m_i"composite families" with demand propensity $\sum_d N_d\, p_d$).

Figure 4 “Families” Community Benchmark versus Market Benchmark, by County, 2019

Figure 4 graphs the Families CB against the MB for each county, using data for 2019. The blue line has a slope of 1.25 to represent the approximate dividing line created by the “best-of clause” and the benchmark multipliers (see Table 4, below). For counties lying above (below) this line, performance against the MB (CB) will be the reference for the “best-of clause”. Equivalently, in counties lying above (below) the line the MB (CB) will be the “easiest-to-beat” benchmark, the metrics for which are provided in the tables included in the Figure.



This is a plausible view of the patterns of CBs and MBs across geographies, which is illustrated in Figure 4. The left panel show that the LI share of families is substantially greater than the LI share of Loans. Indeed, in almost all counties, the MB will be the easiest LI benchmark to beat. For MI families, the picture is more mixed: in some counties the MI share of families exceeds their share of loans and in others it is the reverse. One explanation of these patterns is that LI families’ loan demand is lower than MI families’. LI families may include a higher proportion of people who are not in the market for a loan, because their income is low, and/or because they are very young or very old. MI families may include more prime age adults, who express a relatively higher demand for loans. If potential demand by each LI and MI family is the same, then the only explanation of the LI outcome MB<CB is that lenders are disproportionately rationing loans to LI borrowers.

Table 1 contains four examples. There are two demand groups, A and B, with populations listed in the first three columns. The populations are the same across all four examples, and so the value of the CB is the same in each example: LMI families comprise 31% of all families. Each group has a demand propensity as described above, and lenders have a propensity to extend loans to each group. Because no one can be forced to take a loan, the lending propensity of each group is never greater than its demand propensity. The numerator of the MB is simply the population weighted LMI lending propensity, and the denominator is the same average for the entire population. Last, we include a "(potential) Demand Benchmark" which uses the corresponding population-weighted demand propensities.

Table 1

Demand Group	LMI	Populations Other	Total	Propensities Demand	Lending		Market	Benchmarks Community	Demand
A	200	100	300	0.2	0.2	Numerator	50	400	50
B	200	800	1000	0.05	0.05	Denominator	110	1300	110
Total	400	900	1300			LMI SHARE	0.45	0.31	0.45
A	200	100	300	0.05	0.05	Numerator	50	400	50
B	200	800	1000	0.2	0.2	Denominator	215	1300	215
Total	400	900	1300			LMI SHARE	0.23	0.31	0.23
A	200	100	300	0.05	0.05	Numerator	10	400	50
B	200	800	1000	0.2	0	Denominator	15	1300	215
Total	400	900	1300			LMI SHARE	0.67	0.31	0.23
A	200	100	300	0.05	0	Numerator	40	400	50
B	200	800	1000	0.2	0.2	Denominator	200	1300	215
Total	400	900	1300			LMI SHARE	0.20	0.31	0.23

In the first two cases, each group's demand propensity is the same as its lending propensity. The Market and Demand Benchmarks are thus identical. In the first example, MB>CB. The "best-of" rule would effectively judge banks in this geography against the CB, handing any bank close to the MB a high grade for just meeting potential demand. The second example simply switches the propensities between groups A and B: now the smaller demand group (A) has the smaller demand and lending propensities. The result is MB<CB, and the best-of rule chooses the MB, and would judge close-to-MB banks more strictly.

These two examples tend to contradict the NPR's assertion that "the agencies' proposal would tend to assign better ratings in markets where more banks were meeting the credit needs of the community." (See, e.g., NPR at p.215). Credit needs are met identically in both examples, but the best of clause assigns higher ratings where large groups have small propensities.

The bottom two cases replicate the second case, except banks do not lend to one of the groups. In the third case, the "best-of" clause rewards this behavior by judging banks against the lower CB. In the fourth case, the MB becomes the target of the best-of clause. It is the easiest-to-beat of the three benchmarks, although this may come about precisely because banks restrict loans to one group (for reasons condoned by the CRA or not).

These examples show that the relationship between the MB and CB is dominated by compositional effects, which confound the relationship between the CB and potential demand

and compromise the “best-of clause”. The only reliable thing the Score Test can tell us is how a bank’s LMI share compares with the MB.

2. Secondary components of the RLT account for the vast majority of banks’ failure rates. The design and calibration of these components may prompt unintended consequences from banks, in their efforts to pass the RLT, to the detriment of communities.

The issues described so far are clear from a side-by-side comparison of the RLT and the Lending Test. Our other concerns derive from applying the RLT to the historical CRA data to simulate performance evaluation outcomes and tracing the results –particularly the failure rate—to their sources. While, as the NPR reports, the net failure rate may be in a satisfactory range, it results from combining several features whose design and calibration both are causes for concern.

a) Replicating RLT Results.

We replicated the results of the RLT using data from the CRA data file for 2019. We followed the recipe laid out above, except for the following departures:

- Our Score calculations are only based on Home Loans, and do not include Small Business and Small Farm lending.
- For facility-based AAs we use individual counties for which the variable “County_AA_Flag” is equal to 1. This is because the “Assessment Area Number” variable seemed corrupted (for example, Bank of America only registered two AAs). (For Retail Lending AAs we used the aggregation recipe by state laid out in the NPR.)
- We applied the 60% Test just to a bank’s results for the RLT (restricted to Home Loans), because the data for the other three Tests is not available.
- We measured the loan contribution of an AA to the bank’s aggregate score just by its volume share of Home Loans and exclude Small Farm and Business Loans.

There are 625 banks in the dataset for 2019. Table 2 gives a snapshot of their composition.

Table 2 Shares of Total Banking Assets, Home Loans and Banks by Bank Asset Size (%)
(Figures may not sum to 100 as a result of rounding)

	Bank Assets		
% of	>$50bn	$2bn : $50bn	<$2bn
Assets	79	18	2
Home Loans	58	37	5
Banks	6	52	43

Figure 5: Retail Lending Test Failure Rates

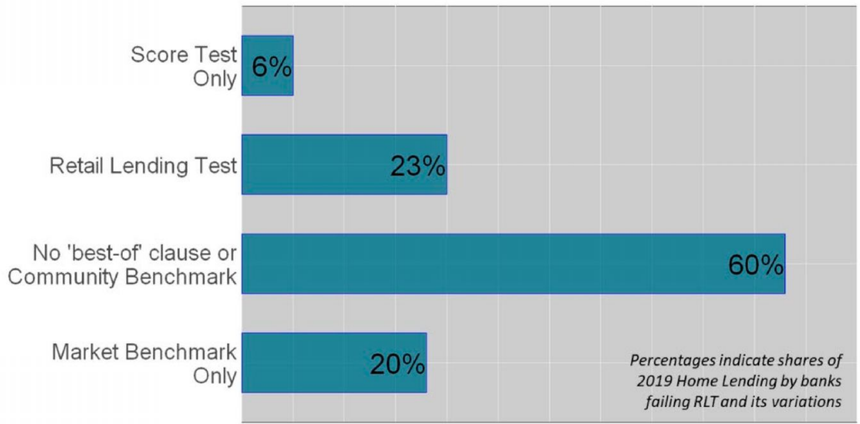


Figure 5 summarizes failure rates for the Retail lending Test and several variations. Banks representing 23% of Home Loan volume would have failed the RLT in 2019, the most recent year in the dataset.

b) <u>The Role of Secondary Features in RLT Outcomes</u>

While the Score Test is the methodological core of the RLT, it only accounts for 6 percentage points of the overall failure rate of 23%. Of the other 17 percentage points, 16 come from the 60% Rule, for which the NPR provides minimal explanation.[15]

We argued above that there is no economic rationale for CBs and the "best-of" clause. When they are eliminated and all other features of the RLT are left the same, the failure rate almost triples from 23% to 60% of Home Loan volume. This version of the Score Test benchmarks banks only against the MB. In 2019, 70% of loans in the MB came from nonbank lenders, whose LI and MI shares were about one-sixth higher than banks. Removing the shield of the CBs exposes banks to direct comparison with these lenders; it pushes the banks over the "cliffs" presented by the Retail Loan Volume Screen and the 60% Test, into failure. Last, when the Score Test is restricted to the MB, and the Retail Loan Volume Screen and the 60% Test are removed, the failure rate drops to 20%, which is very similar to the full Retail Lending Test.

c) <u>Possible Unintended Consequences of RLT Design</u>

The two largest contributors to the overall outcome of the RLT are thus its most questionable individual features. The NPR describes the 60% Rule but **<u>provides no discussion or rationale</u>**. The CB-best-of combination lacks economic foundation, and does not behave as intended, except in very limited and unrealistic cases. This combination of large impact and arbitrariness is fertile ground for unintended consequences for the regulatory effort and communities, as it

15 "The agencies also propose imposing additional restrictions on state, multistate MSA and institution-level ratings for large banks with ten or more assessment areas in a state, a multistate MSA, or overall, respectively. A large bank with ten or more assessment areas (facility-based assessment areas and retail lending assessment areas combined) at the relevant level would not be eligible to receive a "Satisfactory" or higher rating at that level unless it achieved an overall performance of "Low Satisfactory" or better in at least 60 percent of its assessment areas there" (NPR, p.368). Calibrating a test with such impact to a single threshold is likely to be very precarious.

incentivizes banks to cherry pick geographies that are "cheap" in terms of the key metrics, to pass the RLT:

- *Best-of-clause*: banks seek out AAs with low CB values
- *Retail Loan Volume Screen*: banks move out of areas where their lending ratios are low and concentrate where they pass. They may also be disinclined to enter new areas.
- *60% Test:* the simplest way for a bank to pass is, again, to shut down activity in AAs where it is failing, and avoid expansion to risky new areas, with similar consequences for competition.
- While RLVS operates on individual AAs, the 60% Test applies to the bank's overall RLT conclusion, and so the incentive it presents to eliminate failing AAs is that much greater.

To illustrate we examine the strategy of a hypothetical bank with an extreme aversion to making LMI loans. It seeks to minimize the number of LMI loans it makes across all AAs, while passing the RLT. To keep matters simple, we focus on the Borrower perspective only, as there is likely to be significant double counting of the same loans if we aggregate Borrower and Geographic perspectives. The bank allocates its existing total home loans and deposits as well as LMI loans across AAs. As a practical matter, banks of course are not able to allocate their activity freely in this way. The exercise is useful as a way of understanding their incentives.

Say the bank operates in a specific number ("N") AAs, each indexed by *i*, and we denote loan counts by B and deposit amounts (in \$) by D. The bank thus wants to minimize:

$$\sum_{i=1}^{N}(B_{bl,i} + B_{bm,i})$$

subject to:

(1) $\sum_{i=1}^{N} B_i = B, \ \ \sum_{i=1}^{N} D_i = D,$

(2) $Passing\ the\ 60\%\ Test$

(3) *An average score across AAs of no less than 4.5*

(3a) *Each AA score in (3) is weighted by* $0.5 * (\frac{B_i}{B} + \frac{D_i}{D})$.

In words: the bank seeks to allocate its loans, B[16] and deposits, D across AAs, while staying on the right side of the RLT. The RLVS and Score test are embedded in condition (3). Without condition (3a), there would be no material constraint on deposit-taking: if the weights were just AA loan shares, the bank could park all its deposits in one AA and make no loans there. It would fail both RLVS and the score test in the AA, but the weight on the failure in the average score would be zero, and it would be unconstrained by the RLVS in any other AA. The presence of the term $\frac{D_i}{D}$ in (3a) means that deposits affect scores. The larger the deposits in the AA, the larger

16 The subscripts 'bl' and 'bm' stand for bank b's low-income and moderate-income borrowers, respectively.

$\frac{B_i}{B}$ must be to ensure that the RLVS is passed, and then the larger must LMI loans be to ensure the Score test is passed.

While this is a complicated non-standard optimization problem, it is clear that its solution will place a premium on AAs where the following two hurdles are low:

> δ: The ratio of Loans to Deposits for the aggregate of all other banks in the AA. The RLVS requires the bank's Loan-to-Deposit ratio to exceed 30% of this to receive credit for any LMI loans under the score test.
>
> λ: The lower of the MB and CB, scaled by their respective multipliers associated with a LS Conclusion (i.e. 80% and 65%). There will be one value of λ for LI borrowers and one for MI borrowers (λ_{li} and λ_{mi}), which our calculations combine in a simple average.

A low value of δ in an AA means that RLVS is easier to satisfy, making it cheaper to source deposits there, while a low value of λ means that fewer LMI loans are needed to satisfy the Score Test. Neither of these motivations is consistent with the objectives of the CRA or the NPR.

Figure 6 plots δ against λ for counties where more than 5000 home loans were made by all lenders in 2019, and some banks took deposits. It shows substantial variation in the two measures, suggesting that a bank could materially affect its ease of passing the RLT by judiciously selecting where to do business.

Figure 6: Retail Lending Test Hurdles by County, 2019
(See text for definitions. Each circle represents a county)

This analysis is intended as no more than suggestive. It is likely that it is difficult for banks to control independently the loan and deposit magnitudes involved. The data presented in the second section also suggest that banks mainly regulate their lending by controlling total loans, and LMI loan volume adjusts in line. Here we have made the contrary assumption that total loan volume is fixed. Nevertheless, it is worthwhile to understand the incentives the RLT presents.

3. ***A simpler version of the test that excludes the contentious components produces much the same aggregate outcome as the RLT but lessens potentially adverse incentives for banks.***

The contentious components in question are the Community Benchmarks combined with the "best-of clause," which fails to capture potential demand, and the 60% Test, for whose presence and calibration the NPR provides no justification. The first lowers the failure rate dramatically, while the second increases it. As Figure 5 shows, excluding both (and the RLVS) produces a 20% failure rate, which is very similar to the 23% of the full Retail Lending Test.

Table 3 displays failure rates for several combinations of test components and breaks out the aggregate by bank asset size.

Table 3: Failure Rates for Combinations of RLT components (%)
(Figures may not sum to 100 as a result of rounding)

Rule Components	All Banks	Bank Assets >$50bn	Bank Assets $2bn:$50bn	Bank Assets <$2bn
MB, CB	6	2	11	11
MB,CB,RLVS	7	3	13	17
MB, CB, RLVS, 60%	23	14	37	17
MB	20	22	17	23
MB, RLVS	36	45	22	34
MB, RLVS, 60%	60	67	53	34
Likelihood	56	50	66	45

There are two candidates for a simplified test: the Score Test using the MB alone, and the same plus RLVS. The second is desirable because the CRA stresses lending responsibilities of banks where they take deposits. The aggregate failure rate here jumps by 16 percentage points, but this is likely more a matter of the "cliff" nature of the RLVS, rather than of the desirability of including a loan/deposit ratio test.[17] A "smoother" version of the RLVS penalty could have the

[17] It is interesting that when the CB-best-of combination is present, the RLVS has minimal impact, raising the failure rate by only one percentage point (top two bars of Table 3). A failure of RLVS is assumed to reduce the score in an AA to that associated with NI, independent of the bank's performance in the Score Test. In the AAs that pass RLVS, the "MB, CB" Score Test produces a high enough average score that their aggregate with the RLVS failures stays above 4.5 points for almost all banks, and RLVS has only the one percentage point impact. In contrast, the "MB" Score Test produces a lower score for banks even when they pass RLVS, and the aggregate of this group with the RLVS failures pushes 16 percentage points of

desired effect of rewarding adequate lending in relation to deposits, without the discontinuity and interaction effects. We do not have the capacity to explore the numerical consequences of this possibility.

4. The simpler test is directly comparable with a statistical benchmark

The MB test only uses the Market Benchmark. Its outcomes for banks result from the "Multipliers" specified in the NPR, and replicated in Table 4, below.

Table 4: Benchmark Parameters of the Score Test and Likelihood Test

		Benchmark Multipliers (%)				
	Acronym	Market	Community	Score	Community/Market Multiplier Ratio	Likelihood Threshold (%)
Substantial Noncompliance	SN	0	0	0	0.00	0
Needs to Improve	NI	33	33	3	1.00	1
Low Satisfactory	LS	80	65	6	1.23	10
High Satisfactory	HS	110	90	7	1.22	90
Outstanding	OS	125	100	10	1.25	99

We can put these multipliers in context by comparing them with thresholds derived from a more familiar framework: statistical significance. Specifically, we frame the comparison of a bank's LI or MI share with the MB in an AA as a test of a hypothesis:

- *The bank lends to LI and MI borrowers at the same rate as the market*

If this hypothesis is true, the bank's LI or MI share will differ from the corresponding MB only by chance, which can be quantified by a statistical distribution, the hypergeometric distribution. This distribution requires no parameters be estimated to determine the bank's p-value, which measures the likelihood that the bank's lending outcome would be as it is, given that it is lending to LI or MI borrowers at the same rate as the market. We name this test the "Likelihood Test."

To align the Likelihood Test with the NPR scoring rule, we need to map significance levels into "Conclusions." This mapping substitutes for the NPR's "Benchmark Multipliers." Any specific

banks over the 4.5 threshold into failure. Thus, RLVS has a large impact because the CB component is removed. This analysis shows the tight interactions between test components, which might not be intended or desirable.

mapping is of course arbitrary but does have to answer to an external reference: if we set the threshold for SN at 1%, for example, we know that only 1% of banks who conform to the hypothesis will (unluckily) receive a Conclusion of SN. We have no such external reference for the Conclusion cutoffs of the Benchmark Multipliers. The significance cutoffs we use are shown in the last column of Table 4.

Banks' results for the Likelihood Benchmark are shown in the bottom row of Table 3.[18] Likelihood failure rates are materially higher than failure rates for every permutation of RLT components except that which excludes the CB but retains RLVS and the 60% Test. Like the "MB" version of the RLT, the Likelihood test uses a single criterion: performance relative to peers. It is calibrated against a scale of statistical significance, while the "MB" version uses the benchmark multipliers. Results for the Likelihood Test suggest that neither the "MB" nor "MB, RLVS" versions of the RLT are too stringent, as long as one thinks that a 10% statistical outlier warrants failure.

The Likelihood Test failure rates are much higher than one would expect from the luck of the draw. Since we set the threshold for failure at (less than) a 10% likelihood, we would expect 10% of banks to fail even if all banks pursued a strategy in line with the market. ***However, 56% of them fail (as a share of total 2019 loans).***

5. ***We suggest additional features and adjustments that align the new test better with CRA objectives*.**

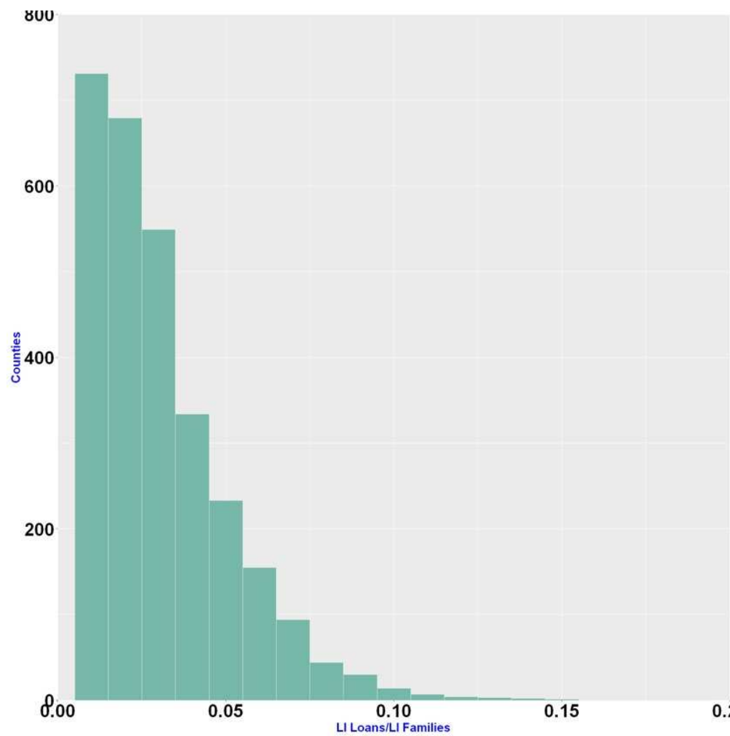
The analysis to this point argues for a stripped-down version of the RLT, which includes tests against the MB in each of the four geography/income group categories in each AA. The RLVS can be added, but a "smooth" as opposed to "cliff" version would be more robust. Comparison with the statistical benchmark provided by the Likelihood Test does not suggest that either of these RLT versions are too stringent.

Two significant issues remain. One is to make the test sensitive to potential demand for loans, which was the intent if not the effect of including CBs and the best-of clause. The other is to make the test responsive to banks exiting home lending and to the presence of redlining.

[18] The version of the Likelihood Test used calculates a statistic for each of the four categories in each of the bank's N AAs and aggregates them into a single number for the bank. The statistic is the standardized deviation of the bank's performance from expected under the hypothesis (both mean and standard deviation depend on known data, and do not need to be estimated). The aggregation uses the same weights as the NPR recipe to combine categories and AAs. The required distribution of the aggregate statistic for the bank can be calculated by Monte Carlo simulation of the 4 * N hypergeometric distributions using the bank and market parameters in each case. 1000 simulations are sufficient for a precise answer. A more accurate version would take into account the dependence among categories in each AA and simulate from N independent multivariate hypergeometrics.

One route to incorporating potential demand is to include a function of aggregate LMI loans transacted per LMI family in the weights used to combine a bank's AA results. Call this variable LPF for convenience. Figure 7 illustrates the range of values for LI LPF across counties.

Figure 7: 2019 LI Loans/Families: Count of Counties



Of course, LPF could vary across geographies for many different reasons, as discussed in section 1c, above. However, one would expect a lot less heterogeneity in LI LPF across counties than in LPF across income groups in the same county. As in that discussion, LPF could be low for demand driven reasons (e.g., very young demographic not in the home loan market), or supply driven reasons (e.g., rationing of loans by lenders). However, if the metric incentivizes LMI lending in low-LPF geographies, banks responding to this incentive, who will have more information than this unavoidably crude metric, will try to expand into the supply-constrained ones.

The version we use to illustrate indeed puts higher weight on counties (AAs) with low LPF. In the current RLT, the weight applied to an AA's composite Score (itself the weighted average of Score Tests of LI and MI borrowers and geographies) is:

$$w_i = 0.5 * (l_i + d_i),$$

where i indexes the AA and l and d are the AA's shares of the bank's total loans and deposits, respectively. Denoting by f the function of LI LPF we use (which depends negatively on LPF), the new weight is:

$$w_i = 0.5 * \left(\frac{l_i f(LPF_i)}{\sum_j l_j f(LPF_j)} + d_i\right).$$

No changes are made to the Score test calculations.

Table 5 repeats Table 3 using the LPF-based weights. The failure rates tend to be higher than in Table 3 for the least stringent tests ("MB,CB," "MB,CB,RLVS," and "MB"), which depend most on the outcomes of the Score Tests. This suggests that banks perform worse on Score Tests where LPF is lowest. In the other cases, which contain the 60% Test, the effect of the Score Test is in effect masked. The changes in results between the two tables are less important than the reassurance that the test is sensitive to potential demand from LI families.

Table 5: RLT Failure Rates with AA weights dependent on LPF (%).
(Figures may not sum to 100 as a result of rounding)

Rule Components	All Banks	Bank Assets		
		>$50bn	$2bn:$50bn	<$2bn
MB, CB	6	2	11	10
MB,CB,RLVS	12	11	14	18
MB, CB, RLVS, 60%	28	22	37	18
MB	28	35	18	21
MB, RLVS	36	46	21	35
MB, RLVS, 60%	60	67	53	35

While incorporating LPF into AA weights allows RLT outcomes to depend on potential demand in a rational manner, it does not deal with the other problem raised at the outset, which is that the RLT's dependence on static measures of shares means that it cannot detect uniform exit from lending by banks, or uniform restrictive practices such as redlining. It would be desirable to incorporate a variable that monitored the rate of change of bank lending, but we have not identified a way of doing this that does not run up against scaling problems. Instead, it seems best to approach this from the standpoint that the absolute level of lending is important, and the appropriate scale variable here seems to be deposits. The RLVS component of the RLT is not up to this task, because it benchmarks a bank's loan/deposit ratio against those of other banks in the same geography. If all banks reduce loans in unison, RLVS will not sound an alarm and the banks will earn a passing score. The CRA was built on the idea that a bank that took deposits in a geography had an obligation under its charter, and because it benefitted from deposit insurance, to provide other banking services in that geography, including loans. One way of looking at this is that deposit insurance premiums are partly paid in kind by banks, by making loans. We suggest that CRA evaluations recognize this explicitly, and put a loan price on deposits, e.g., annual loan origination value in a geography has to exceed 10% of average annual deposits. Unless a bank is willing to reduce its deposits, it cannot reduce its loans without penalty. And as long as it makes loans in general, the score test will require the bank to make LMI loans in line with other lenders. This will not avoid redlining. However, if the loan percentage of deposits requirement is income-based, disincentives to redline will result.

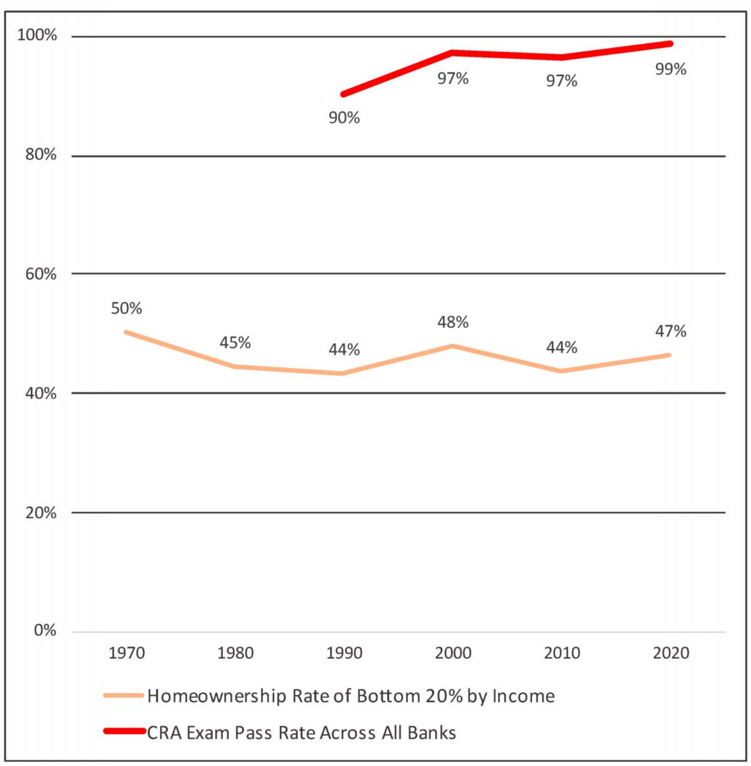
Conclusion

Our analysis – the only detailed statistical analysis of the rule proposed in the NPR – shows that the proposed rule will **not** work as intended and will **not** detect textbook cases of redlining, but it will amplify indefensible deficiencies of the current rule. It, therefore, will not achieve the stated goals set forth in the NPR or the mandate and intent of the CRA itself.

As we said above, redlining has inflicted egregious harm on LMI communities for many decades. By constricting the flow of capital to LMI communities, redlining has created, exacerbated, and perpetuated inequality, poverty, and racism, often intergenerationally. The result has been to prevent investment, equity and wealth building in LMI communities, which has ignited a downward spiral that has destroyed too many lives, neighborhoods, and dreams for far too long.

Two data points illustrating the failures of the current CRA rule are reflected in the Figure 8 below. It shows that homeownership among the bottom 20% by income (a proxy for LMI communities) has actually declined in the 45 years since the CRA was enacted into law. To highlight how poorly the current CRA rule has performed during that time, as homeownership was declining, the CRA exam pass rate across all banks has gone up from 90% in 1990 to 99% in 2020:

Figure 8: The Homeownership Rate Decline Among Low-Income Families is Inconsistent with the Nearly 100% CRA Pass Rate for Banks



Source: U.S. Census Bureau Decennial Surveys

While these are just two data points, numerous others are detailed above. Unfortunately, the likely outcome of the proposed CRA rule is equally bleak if the NPR is finalized roughly as proposed. Only by making material changes as suggested herein might the future not replicate the failures of the past.

We hope these comments are helpful and we would be pleased to discuss them with you further.

Sincerely,

Dennis M. Kelleher
Co-Founder, President and CEO

Peter Rappoport
Senior Fellow

Better Markets, Inc.
2000 Pennsylvania Avenue, NW
Suite 4008
Washington, DC 20006
(202) 618-6464

CC:
Ann E. Misback
Secretary
Board of Governors of the Federal Reserve System
20th Street and Constitution Avenue NW
Washington, DC 20551

Chief Counsel's Office
Attention: Comment Processing
Office of the Comptroller of the Currency
400 7th Street SW, Suite 3E-218
Washington, DC 20219

James P. Sheesley
Assistant Executive Secretary
Attention: Comments RIN 3064-AF81
Federal Deposit Insurance Corporation
550 17th Street NW
Washington, DC 20429

Eric S. Belsky, Director, Division of Consumer and Community Affairs, *Fed*
Joseph A. Firschein, Associate Director, Division of Consumer and Community Affairs, *Fed*
S. Caroline Johnson, Manager, Division of Consumer and Community Affairs, *Fed*
Amal S. Patel, Counsel, Division of Consumer and Community Affairs, *Fed*
Christine Graham, Special Advisor, *Fed*
Laura Lipscomb, Special Advisor, *Fed*
Jonathan Miller, Deputy Director of Consumer Policy & Research, *FDIC*
Patience R. Singleton, Senior Policy Analyst, Supervisory Policy Branch, Division of Depositor and Consumer Protection, *FDIC*
Pamela Freeman, Chief Fair Lending and CRA Examination Section, Division of Depositor and Consumer Protection, *FDIC*
Richard M. Schwartz, Counsel, Legal Division, *FDIC*
Kristopher Rengert, Senior Consumer Researcher, *FDIC*
Heidi Thomas, Special Counsel, *OCC*
Emily Boyes, Counsel, Chief Counsel's Office, *OCC*
Vonda Eanes, Director for CRA and Fair Lending Policy, *OCC*
Karen Bellesi, Director for Community Development, Bank Supervision Policy, *OCC*