URBAN INSTITUTE · ELEVATE · THE · DEBATE

Tuesday, October 15, 2019

Consumer Financial Protection Bureau

1700 G Street NW

Washington, DC 20552

RE: Advance Notice of Proposed Rulemaking Relating to Home Mortgage Disclosure (Regulation C) Data Points and Coverage, Docket Number: CFPB-2019-0020; RIN 3170-AA97

By electronic submission to http://www.regulations.gov

Dear Director Kraninger:

Thank you for this opportunity to comment on the Consumer Financial Protection Bureau's advance notice of proposed rulemaking (ANPR) raising the prospect of substantially amending Regulation C to decrease information available to regulators and the public under the Home Mortgage Disclosure Act (HMDA). We are researchers in the Housing Finance Policy Center at the Urban Institute, and we frequently use HMDA data.[1] As is Urban's policy, these comments are our own and should not be attributed to the Urban Institute, its trustees, or its funders.

The first part of this millennium was marked by a major housing market boom and bust. Although there are still arguments about the causes, we know that regulators missed much of what was happening in the housing market. As late as May 2007, Federal Reserve chairman Ben Bernanke said, "We believe the effect of the troubles in the subprime sector on the broader housing market will be limited, and we do not expect significant spillovers from the subprime market to the rest of the economy or to the financial system."[2]

One of the reasons the Federal Reserve and other regulators—as well as other policymakers in and outside the government—missed the warning signs was that they did not have sufficient standardized data about individual loans in the huge and rapidly changing mortgage market. Regulators and policymakers looked at what they had collected through HMDA, but those data were missing critical items that could have led policymakers to question their assumptions about the extent of layered risk in and beyond the subprime market.

With the Dodd-Frank Act in 2010, Congress, responding to the crisis, created the Consumer Financial Protection Bureau (CFPB) and charged it with revising HMDA regulations to include 13 new data points. These data points were largely related to issues that regulators had long been unable to assess (e.g., credit

---

[1] Our biographies can be found at https://www.urban.org/policy-centers/housing-finance-policy-center/researchers.

[2] Ben S. Bernanke, "The Subprime Mortgage Market" (speech, Board of Governors of the Federal Reserve Board Chicago, IL, May 17, 2007), https://www.federalreserve.gov/newsevents/speech/bernanke20070517a.htm.

scores) or that would have clearly signaled risk layering (e.g., nonamortizing features, introductory rate periods, prepayment periods, and points and fees). In addition, the CFPB was given the authority to add "such other information as the Bureau may require."[3]

In August 2014, the CFPB published a proposed regulation implementing section 1094 of Dodd-Frank. The proposal filled 151 Federal Register pages, laying out the CFPB's proposals on the required Dodd-Frank elements—including modifications and limitations meant to balance the public interest and regulatory burden—and on additional elements the CFPB determined necessary to effectuate HMDA's purposes, namely "(i) to help determine whether financial institutions are serving the needs of their communities; (ii) to assist public officials in distributing public-sector investment so as to attract private investment to areas where it is needed; and (iii) to assist in identifying possible discriminatory lending patterns and enforcing antidiscrimination statutes."[4] Among the elements the CFPB added were further signs of risk layering, such as combined loan-to-value ratios (related to piggyback loans), debt-to-income ratios, lender credits, open-end mortgages, and more detailed information on reverse mortgages and two market sectors critical to housing affordability: multifamily and manufactured housing loans. In 2014, we and our Urban Institute colleagues submitted extensive comments on the CFPB's comprehensive proposed update of Regulation C.[5]

On October 28, 2015, the CFPB released amendments to Regulation C implementing changes to HMDA (the "2015 rule"). As did the initial proposal, the final regulation thoroughly assessed the regulation's benefits and burdens in the context of HMDA's critical role as a disclosure statute for both regulators and the public. The new data requirements largely came into effect on January 1, 2018.

In May 2019, the CFPB released an ANPR that opens the possibility of substantially modifying or eliminating some of the data added to the regulation in response to the issues that led to the housing crisis. The ANPR comment period was initially a short 60 days. It was extended to allow people to evaluate the 2018 HMDA data, including new data elements. Those data were released in August 2019.

We understand that the ANPR aims to better balance the costs of collecting data with the benefits that can accrue from the data. We agree that a balance is required but are concerned that the current ANPR could reverse many of the decisions made in 2014 and 2015. We are particularly concerned because the bureau is raising this possibility even though we are only beginning to be able to evaluate data newly available pursuant to the 2015 rule. Moreover, given that reporting entities have changed their systems to enable them to comply with the 2015 rule, further changes are likely to increase regulatory burden rather than mitigate it.

The CFPB asked about four pressure points in reporting: manufactured housing, disaggregated race or ethnicity, open-format fields for loan denial, and multifamily loans to corporate entities. We will review each. Then, we will turn to some of the benefits from the enhanced reporting. We should note, too, that the CFPB released two excellent reports on the 2018 HMDA data in August 2019: *Data Point: 2018 Mortgage Market Activity and Trends* (hereafter, the "market report") and *Introducing New and Revised Data Points in HMDA* (hereafter, the "new data report"). These reports make an excellent case for the value of the additional data.[6]

---

3 Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, 124 Stat. 1376-2223 (2010), section 1094.

4 Authority, Purpose, and Scope, 76 Fed. Reg. 78468 (December 19, 2011).

5 Ellen Seidman, Laurie Goodman, Wei Li, Carlos Martín, Jim Parrott, Kathryn L.S. Pettit, and Peter A. Tatian, "Comment Letter on the CFPB's HMDA Mortgage Data Proposal (Washington, DC: Urban Institute, 2014).

6 Consumer Financial Protection Bureau (CFPB), *Data Point: 2018 Mortgage Market Activity and Trends, A First Look at the 2018 HMDA Data* (Washington, DC: CFPB, 2019); and CFPB, *Introducing New and Revised Data Points in HMDA: Initial Observations from New and Revised Data Points in 2018 HMDA* (Washington, DC: CFPB, 2019).

Manufactured Housing

Manufactured housing differs from site-built housing in several ways, one of the most important being that the homeowner quite frequently does not own the land on which the home sits. The CFPB stated its concern that it is more burdensome to report information about whether a borrower owns or leases the land on which a manufactured home sits, as the information is not generally collected in the normal course of business. Yet this information is important as we try to address the lack of affordable housing. A recent presidential memorandum is testimony to the importance of this issue.[7] And manufactured housing is the most affordable housing, with a cost per square foot well under that of site-built housing.[8]

But one of the obstacles to increasing the manufactured housing supply is that financing is expensive. Borrowers can get a traditional mortgage, but only if the lien is on the structure and the land. If a borrower does not own the land or chooses not to encumber it, he or she can obtain a personal property or chattel loan only on the structure, which is more expensive.

As the nation looks to create more affordable housing, we need to know how many manufactured housing owners who also own their land nevertheless take out a personal property loan. Before 2018, we could only guess, but the new 2018 HMDA data provide that information.

For manufactured housing purchase loans, there were 77,544 loans on which the borrower owned the land free and clear. But only 62,771 of these loans were secured by both the manufactured home and the land, while 14,691 borrowers took out a chattel loan when they might have been able to take out a real estate loan.[9] Table 1 shows the characteristics of these borrowers.

---

7 "President Donald J. Trump is Tearing Down Red Tape to Build More Affordable Housing," White House, June 25, 2019, https://www.whitehouse.gov/briefings-statements/president-donald-j-trump-tearing-red-tape-order-build-affordable-housing/.

8 Laurie Goodman, Edward Golding, Alanna McCargo, and Bhargavi Ganesh, "Manufactured Homes Could Ease the Affordable Housing Crisis. So Why Are So Few Being Made?" *Urban Wire* (blog), Urban Institute, January 29, 2018, https://www.urban.org/urban-wire/manufactured-homes-could-ease-affordable-housing-crisis-so-why-are-so-few-being-made.

9 This split shows a higher share of real estate loans versus chattel loans for borrowers who owned their own land than previous estimates. See Consumer Financial Protection Bureau, *Manufactured-Housing Consumer Finance in the United States* (Washington, DC: Consumer Financial Protection Bureau, 2014); and Laurie Goodman and Bhargavi Ganesh, "Challenges to Obtaining Manufactured Home Financing" (Washington, DC: Urban Institute, 2018).

TABLE 1

Characteristics of Borrowers Who Have Direct Ownership of Their Land, by Loan Type

Secured property	Number of loans	Median loan amount ($)	Median income ($)	Median interest rate (%)	Median DTI ratio (%)	Median LTV ratio (%)	Black borrower share (%)	Hispanic borrower share (%)
Manufactured home and land	62,771	115,000	52,000	5.25	38	95.53	2.85%	9.48%
Manufactured home and NOT land	14,691	65,000	54,000	7.99	33	88.40	6.68%	12.97%
Not applicable	82	205,000	70,000	5.00	42	95.00	0.00%	7.32%

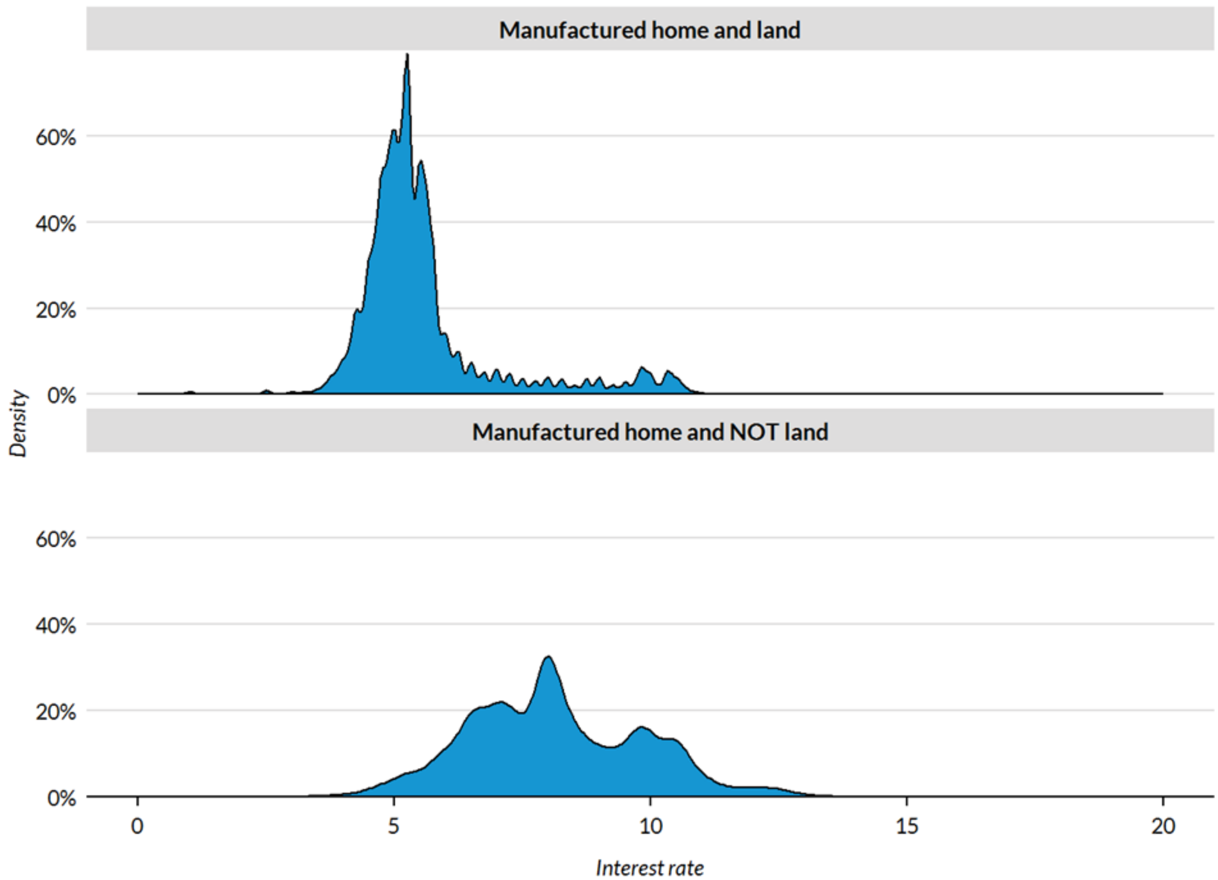
Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Note: DTI = debt-to-income; LTV = loan-to-value.

The average annual income among manufactured home borrowers who owned the land free and clear and took out a real estate loan was $52,000, little different from the income of borrowers who took out a chattel loan ($54,000). The average real estate loan amount was $115,000, substantially larger than the $65,000 for chattel loans. Part of the differential is the cost of the land. The most critical information the new data provide is that the median interest rate is higher for chattel loans (7.99 percent) than for real estate loans (5.25 percent). Figure 1 shows the distribution of these interest rates. Not only is the average higher, but there is a higher share of chattel loans with interest rates significantly higher than the average. The data also demonstrate that black and Hispanic borrowers took out chattel loans at greater rates than real estate loans. We do not know whether lower credit scores might have driven borrowers (of any race) to take out chattel loans. Credit score information was collected and reported pursuant to the 2015 rule, but the CFPB has not released it, even in buckets—a loss of beneficial information without any reduction in lender burden.

FIGURE 1

Distribution of Interest Rates for Borrowers Who Have Direct Ownership of Their Land, by Mortgage Type



Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Fannie Mae and Freddie Mac are expanding their efforts to ease financing issues for manufactured housing as part of their Duty to Serve mandates.[10] These new data could inform these efforts. In particular, it is vital to understand (1) the characteristics of borrowers taking out chattel loans when they could be taking out real estate loans and (2) the difference in interest rates and costs between the two loan types. To increase the supply of affordable manufactured housing, we need to be able to assess whether a borrower can obtain a real estate loan. This is impossible without knowing whether the borrower has direct ownership of the land. In short, we need the data that the CFPB is considering eliminating.

Disaggregated Race or Ethnicity

Before this latest set of data, financial institutions were required to ask loan applicants their race or ethnicity using aggregated categories for race (e.g., white, black, Asian) and ethnicity (Hispanic or Latino, not Hispanic or Latino). Beginning in 2018, HMDA requires the lender to identify race using disaggregated

[10] Fannie Mae, "Duty to Serve Underserved Markets Plan for the Manufactured Housing Market" (Washington, DC: Fannie Mae, 2019); and Freddie Mac, "Freddie Mac's Duty to Serve Underserved Markets Plan for 2018–2020" (Tysons Corner, VA: Freddie Mac, 2018).

categories (e.g., Chinese or Korean) and ethnicity using disaggregated categories (e.g., Mexican, Cuban) in addition to the aggregated categories.

Of the five base racial categories (white, black or African American, American Indian or Alaska Native, Asian, and Native Hawaiian or other Pacific Islander), detailed choices were available for Asian and Native Hawaiian and other Pacific Islander. Our analysis of the data presented in the CFPB's new data report indicates that for the Asian category, 56.9 percent of the loans contained an expanded race (397,800 of 699,600 loans).[11] But some lenders reported expanded race as the first variable. For example, instead of listing Asian as the base racial category and Chinese as the second, leaders listed Chinese as the first, with no second. If we include these additional 97,200 loans, we find that 62.1 percent of the loans (495,000 of 796,800) contained the expanded information.

For the Hispanic or Latino ethnicity, 41 percent contained an expanded race response. But because 7.8 percent of the total selected "other Hispanic or Latino," which provides little information, we have detailed information on 33.2 percent of the total. However, 118,400 selected an expanded ethnicity, such as Mexican, as the first ethnicity. If we include these, 45.8 percent of the loans provided expanded ethnicity information, although 10.8 percent selected "other Hispanic or Latino."

These results show the efficacy of the new reporting categories. Moreover, reporting levels may rise in the future, as lenders gain experience with the question. But as financial technology (fintech) becomes more common in mortgage origination, the lender is not going to be able to observe the borrower's race or ethnicity, and the borrower has little reason to provide it, potentially resulting in a decline in all race and ethnicity reporting, as the CFPB's market report found happened between 2017 and 2018.[12] We think it would be premature to change this data requirement, though more explicit direction on how to answer the question is required.

Open Format for Denial Rates

The 2015 rule requires lenders to report the reason for denying a loan. If none of the standard reasons fit, the lender must report the reason for loan denial as "other" and complete an open-format description of the reason. Because of stakeholder feedback, the CFPB has proposed to eliminate this open-format description.

Table 2 shows the number of reasons for denial for forward loans. Out of the 1.7 million forward loan applications that were denied, 81.1 percent listed one reason for denial, 16.1 percent listed two reasons, and the balance listed three or more. The CFPB notes in its market report, "So, unlike 2017 when only 72.1 percent of denied home-purchase applications and 66.2 percent of denial refinance applications had at least one reported denial reason, all denials in the 2018 HMDA data included denial data."[13]

[11] CFPB, *Introducing New and Revised Data*, 103, 106, 108, and 110.

[12] CFPB, *Data Point*, 22.

[13] CFPB, *Data Point*, 41.

TABLE 2

Number of Reasons for Denial of Forward Mortgages

Number of reasons given	Number of loans
1	1,370,930
2	272,190
3	42,044
4	4,515
Total forward mortgage denials	1,689,679

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Table 3 shows the reasons for denial. Credit history is the most common reason for denial, followed closely by debt-to-income ratio. Collateral, "other," and credit application incomplete are the next three most common reasons.

TABLE 3

Top Reasons for Denial of Forward Mortgages

Reason for denial	Number of applications
Credit history	516,877
Debt-to-income ratio	507,883
Collateral	293,981
Other	235,765
Credit application incomplete	229,319
Unverifiable information	108,083
Insufficient cash (down payment closing costs)	86,406
Employment history	43,510
Exempt	35,003
Mortgage insurance denied	2,675
Total forward mortgage denials	1,689,679

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

The "other" category was used for the 235,765 loans (about 14 percent of loans) for which a denial reason was reported. This is a significant share, suggesting the open-format field is valuable. The CFPB could reduce the number of instances where lenders must use the open-format field by expanding the list of reasons for denial. This can be done by analyzing 2018 HMDA filings and including the most common free-form text responses as choices. In addition, providing more guidance on what should be included in some of the provided choices such as "collateral" or "debt-to-income ratio" would transfer some of the reasons reported as "other" to one of the standard categories.

Curtailing Data on Multifamily Lending

The ANPR asks for comments about whether the CFPB should exclude coverage of business or commercial loans made to a nonnatural person (i.e., a business entity) and secured by a multifamily dwelling. The bureau asserts these data are unnecessary to fulfill HMDA's purpose and that the benefits of reporting do not outweigh the costs. We cannot think of anything more central to HMDA's purpose than multifamily lending, mostly to support the rental market, the country's largest source of affordable housing. This is the most damaging of the ANPR's suggestions, as it would gut multifamily reporting.

For this comment letter, *multifamily dwellings* are those that have five or more units, consistent with the definitions the Federal Housing Administration (FHA), Fannie Mae, and Freddie Mac use.

How many of these loans are to nonnatural persons? The 2018 HMDA data do not tell us directly, but we can infer the count from loans missing data that are not required or available for nonnatural persons, namely income, debt-to-income (DTI) ratio, race or ethnicity, age, or gender. Table 4 shows the number of loans in multifamily structures that are missing these variables. Of the 51,380 multifamily loans, only 13,632 have completed income or DTI ratios, while 37,748 are missing income and DTI ratios. Most loans in both categories are also missing age, race or ethnicity, and sex. We use two different definitions of nonnatural person. The broad definition is loans missing income and DTI ratios (37,748 of 51,380, or 73.4 percent), and the narrow definition is loans missing income, DTI ratios, race or ethnicity, age, and sex (33,729 of 51,380, or 65.6 percent). The overwhelming majority of multifamily loans are to nonnatural persons.

TABLE 4

Missing Variables in Multifamily Loans

	Number of loans	Missing race or ethnicity	Missing age	Missing sex
Completed income or DTI ratio	13,632	9,993	9,693	9,611
Missing income and DTI ratio	37,748	34,908	34,324	33,815

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Note: DTI = debt-to-income.

If we break this down by structure type, we can see that the simple loan count will understate the number of units for which we would lose information (table 5). The table shows that for structures with 5 to 24 units, using the broad definition, 84.2 percent of the units secure loans to nonnatural persons. And this number rises as the number of units increases. For structures with more than 149 units, 92.9 percent of the units secure loans to nonnatural persons. Eliminating this information would render the multifamily part of HMDA useless.

TABLE 5

Nonnatural Borrowers by Number of Units in Structure

Units in structure	Number of loans	Number of loans: NN broad	Number of loans: NN narrow	Number of units	Number of units: NN broad	Number of units: NN narrow	NN broad share	NN narrow share
5–24	32,712	21,924	18,408	474,324	317,898	266,916	67.02%	56.27%
25–49	6,601	5,171	4,882	237,636	186,156	175,752	78.34%	73.96%
50–99	4,864	4,060	3,942	362,368	302,470	293,679	83.47%	81.04%
100–149	2,243	1,987	1,956	279,254	247,382	243,522	88.59%	87.20%
> 149	4,960	4,606	4,541	992,000	921,200	908,200	92.86%	91.55%
All, by number of loans	51,380	37,748	33,729	2,345,582	1,975,106	1,888,069	73.47%	65.65%
All, by number of units	51,380	37,748	33,729	2,345,582	1,975,106	1,888,069	84.21%	80.49%

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Note: NN = nonnatural.

The information loss would be severe. Although banks collect these data for the Community Reinvestment Act and other regulatory purposes, the lack of HMDA data would make it more difficult for examiners to see how one institution compares with similar institutions. The public would lack data on multifamily lending to low- and moderate-income (LMI) areas in the aggregate and total LMI lending to individual areas and would not be able to view the lending patterns of individual institutions or compare institutions as to their relative activity providing loans to LMI areas. In short, the public would lose all visibility on multifamily lending.

In addition, if loans to nonnatural persons were eliminated, the affordability data would be too slim to be useful. As it is, there is a fair amount of nonreporting on affordability. Only 24,284 of the 51,380 loans (47 percent) provide information on affordability. But when the affordability is reported, it is disproportionately on loans to nonnatural persons. The number of fields left blank suggests this is an issue the CFPB will want to address, but eliminating the field directly or eliminating its utility is not the solution to the missing information.

TABLE 6

Affordability for Nonnatural Borrowers

Number of units	Number of loans	Number of loans with affordable units	Share of loans with 100% affordable units	Share of structures with 100% affordable units to NN borrowers (narrow)	Share of structures with any affordable units to NN borrowers (narrow)
5–24	13,863	1,251	5.19%	62.31%	70.18%
25–49	3,327	582	8.12%	90.74%	92.44%
50–99	2,684	690	14.42%	98.45%	98.55%
100–149	1,360	387	15.07%	99.02%	97.42%
> 149	3,050	656	11.34%	99.42%	99.39%
All	24,284	3,566	7.94%	84.12%	87.63%

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.

Note: NN = nonnatural.

Other Data

The new data required by the 2015 rule are valuable, and we applaud the CFPB for allowing most of them to be made public. In this section, we detail some of the uses of these data. Though not comprehensive, this section shows how the new data fields benefit current research and how they can help provide early warning of potential problems in the housing finance market.

First, the new data fields allow regulators, policymakers, researchers, advocates, and the media to monitor if, when, and to what extent the risky product types that contributed to the financial crisis emerge. The CFPB new data report shows these product types are not an issue at present.[14] Only 2.1 percent of loans have balloon payments, 2.3 percent have interest-only payments, and virtually none have negative amortization or other nonamortizing features. Urban Institute analysis shows that virtually no government or government-sponsored enterprise (GSE) loans have balloon or interest-only features. These features are confined to the bank portfolio and private-label securities markets. We found that if the data are cut by loan balances, 5.7 percent of conventional non-GSE loans have interest-only payments. This reflects the fact that interest-only loans are more prevalent in jumbo loans; 14.3 percent of jumbo loans have this feature.

Nevertheless, the 2018 data provide a valuable baseline. Subsequent HMDA data will allow us to monitor the reemergence of risky product types. The data also allow us to examine the differential between combined LTV ratios and LTV ratios. This differential reflects the presence of second mortgages such as piggyback loans, which substantially increase first mortgage risk. Curtailing any of this risk-warning data would put regulators and the public back in the data-blind position they were in before the housing crisis.

The new fields also allow us to look at average loan costs (origination costs and charges for other services related to closing). The CFPB new data report shows that the costs of originating government loans are higher than for conventional loans.[15] That is, total loan costs are $3,357 for conforming conventional

[14] CFPB, *Introducing New and Revised Data*, 163–64.

[15] CFPB, *Introducing New and Revised Data*, 232.

loans and $5,394 for jumbo conventional loans, versus $6,868 for FHA loans and $5,522 for US Department of Veterans Affairs (VA) loans. Most of this is likely because of the 1.75 percent up-front FHA mortgage insurance premium and the 2.15 percent up-front VA funding fee, but the higher cost of government lending is a topic that would suggest further study, as government borrowers are disproportionately first-time homebuyers and minority homebuyers. Moreover, the CFPB new data report shows that black and Hispanic white borrowers pay more in total loan costs than non-Hispanic white borrowers, regardless of channel.[16]

HMDA also contains information on discount points and lender credits, and the rate spread variable is included for all loans. Before 2018, the rate spread (the difference between the annual percentage rate and the average prime offer rate) was disclosed only when the rate spread was more than 200 basis points. The additional disclosure is useful for many purposes and is particularly useful as the qualified mortgage rule is evaluated. A 150 basis-point rate spread is the cutoff between safe harbor and rebuttable presumption under the rule. One fear was that there would be a clustering just under the cutoff. The new data show this is not the case, an important observation that would have been unavailable absent the new HMDA data fields.[17] Similarly, the new data on the DTI ratio, when analyzed by race or ethnicity, was valuable when we provided input to the CFPB's reconsideration of the qualified mortgage rule. Using these data, we could show definitively that black and Hispanic borrowers were more apt to have high DTI ratios, an important point in our qualified mortgage comment letter (table 7 is reproduced from that comment letter).[18]

16 CFPB, *Introducing New and Revised Data*, 234.

17 CFPB, *Data Point*, 51.

18 Karan Kaul, Laurie Goodman, and Jun Zhu, "Comment Letter to the Consumer Financial Protection Bureau on the Qualified Mortgage Rule" (Washington, DC: Urban Institute, 2019).

TABLE 7

High-DTI Lending by Channel and Race or Ethnicity, 2018

Lending channel	Race or ethnicity	Share of loans with DTI ratios > 45%	Number of loans with DTI ratios > 45%	Distribution of all loans	Distribution of all loans with DTI ratios > 45%
Government	Black	43.3%	88,307	13.7%	16.1%
	Hispanic	46.8%	112,854	16.2%	20.6%
	White	32.9%	324,370	66.5%	59.2%
	Asian	47.9%	19,291	2.7%	3.5%
	Other	28.0%	3,526	0.8%	0.6%
	Total	36.9%	548,349	100.0%	100.0%
GSE conventional	Black	21.3%	16,624	4.8%	5.5%
	Hispanic	24.2%	39,702	10.0%	13.1%
	White	17.2%	218,049	77.7%	71.9%
	Asian	24.0%	27,615	7.0%	9.1%
	Other	18.3%	1,394	0.5%	0.5%
	Total	18.6%	303,385	100.0%	100.0%
Non-GSE conventional	Black	18.4%	19,786	5.1%	5.2%
	Hispanic	21.5%	43,825	9.7%	11.4%
	White	17.9%	292,220	77.4%	76.2%
	Asian	16.5%	25,646	7.4%	6.7%
	Other	21.3%	1,911	0.4%	0.5%
	Total	18.2%	383,387	100.0%	100.0%

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.
Notes: DTI = debt-to-income; GSE = government-sponsored enterprise. Includes owner-occupied, single-family (one to four units), site-built, purchase, and refinance loans. Loans with missing race or ethnicity were reallocated based on the share of lending to each race or ethnicity.

Age is also a new data field. Table 8 shows age by channel. One surprising result is the age distribution of rural housing loans. Even though the rural population is older, Rural Housing Service and Farm Services Administration loans skew the lowest of any channel. When looking at who is served by this and other government programs, age is critical information.

TABLE 8

Borrower Age by Channel

Borrower age	Conventional	FHA	RHS/FSA	VA
< 25	3.29%	6.59%	19.11%	3.12%
25–34	24.22%	30.81%	42.88%	23.38%
35–44	24.49%	27.54%	19.58%	21.49%
45–54	20.71%	20.01%	10.62%	19.41%
55–64	16.14%	10.60%	5.52%	13.91%
65–74	8.54%	3.59%	1.88%	14.39%
> 74	2.48%	0.81%	0.38%	4.26%

Source: Authors' calculations using 2018 Home Mortgage Disclosure Act data.
Note: FHA = Federal Housing Administration; RHS/FSA = Rural Housing Service/Farm Services Administration; VA = US Department of Veteran Affairs.

The 2018 age data are also useful to counteract incorrect impressions. For example, many people have asserted that young applicants are more apt to be turned down for mortgages. And when they get mortgages, their DTI ratios are higher because of student loan debt. In fact, CFPB and Urban Institute analyses of the data show that DTI ratios are similar, regardless of age. And the CFPB analysis shows younger applicants are less likely to be turned down for mortgages than older applicants.

Data Disclosure

There are several areas where data are collected, but the CFPB has not disclosed them. Enhanced disclosure, however, would be valuable to achieving HMDA's purposes. Also, because the CFPB, not lenders, does the disclosure, nondisclosure does not decrease lender burden.

A prime example is credit score data, which are collected and reported but not disclosed. We would like to see these data released even if in broad buckets, allowing researchers, nonregulatory policymakers, and the public to analyze data on borrowers by age, income, DTI ratio, LTV ratio, credit score, and race or ethnicity. This would allow us to more vigorously examine the possibility of discrimination by race or ethnicity, income, or geography. In fact, the CFPB included credit score data in several tables and charts in its summary of the data, and they are illuminating. For example, the denial rates for black borrowers and Hispanic white borrowers are significantly higher than for non-Hispanic white borrowers. The CFPB notes in its market report, "Previous research and experience in the fair lending supervisory process show that differences in denial rates and in the incidence of higher-priced lending…among racial and ethnic groups stem, at least in part, from factors related to credit risk. Those factors—such as credit history (including credit score), ratio of total monthly debt to total monthly income (DTI ratio), and CLTV ratio—are now revealed for the first time in the 2018 HMDA data."[19]

The CFPB new data report shows that black and Hispanic borrowers have lower credit scores. The average credit score for black borrowers was 691 but was 710 for Hispanic white borrowers and 748 for non-Hispanic white borrowers.[20] Disclosing credit scores, even in broad buckets, would allow for more conclusive research. Moreover, as fintech lenders become more prevalent, credit score data would allow us to see the full set of characteristics of fintech borrowers. That is, fintech appears to have increased mortgage origination efficiency but does not appear to have broken down structural barriers as it has in other markets.[21] The ability to monitor whether the fintech borrower base is widening will become increasingly valuable.

For multifamily loans, it would be beneficial to disclose the actual number of multifamily units securing the loan and the number of income-restricted units, rather than total unit information grouped into large ranges (e.g., 5 to 24 units, 25 to 49 units) and the share of income-restricted units as a share of total units. This would give us a fuller picture of lending for new income-restricted units. Policymakers and the public would be able to use HMDA data to understand the loan amount per unit (and thus potentially rents) in properties with new multifamily loans. It would also make it possible for people using public data to determine how many income-restricted units a property has. This is valuable information for people outside regulatory agencies who want to understand how well a bank is servicing its community.

Conclusion

The new data fields in the 2018 HMDA data represent a huge step forward in transparency. These data form an important baseline with which to monitor lending conditions. It also gives us more detail on lending by race or ethnicity and age, as well as points and fees. The CFPB's own reports show how much the new data fields contribute to our understanding of the market.

[19] CFPB, *Data Points*, 40–41.

[20] CFPB, *Introducing New and Revised Data*, 195.

[21] Jung Hyun Choi, Karan Kaul, and Laurie Goodman, "Fintech Innovation in the Home Purchase and Financing Market: Impact and Gaps" (Washington, DC: Urban Institute, 2019).

We have made several recommendations in this comment letter for increased disclosure: the disclosure of credit scores by bucket and, in multifamily lending, disclosure of the number of units in each structure and the number of income-restricted units. Enhancing this disclosure does not increase lender burden.

The CFPB asked for feedback on four items where they were concerned that the costs of data collection might outweigh the benefits: land ownership in manufactured housing, disaggregated race or ethnicity, open-format fields in loan denial, and the reporting of multifamily loans to corporate entities. We have shown the importance of understanding land ownership in manufactured housing. Loan denial reasons are important. We have argued that burden can be reduced by incorporating more reasons and better defining current reasons. If reporting of multifamily loans to corporate entities were suspended, it would render the multifamily part of HMDA useless. Although the use of disaggregated race or ethnicity is not as strong on a national level as one might hope, the data themselves suggest that clearer instructions might increase usage and thus the data's utility.

We appreciate the opportunity to comment on this critical data source.

Sincerely,

Laurie Goodman
Vice President
Housing Finance Policy Center
Urban Institute
lgoodman@urban.org

Sarah Strochak
Research Analyst
Housing Finance Policy Center
Urban Institute
sstrochak@urban.org

Ellen Seidman
Nonresident Fellow
Housing Finance Policy Center
Urban Institute
eseidman@urban.org