

Reshape the Design of HMDA Data to Capture the Scope of Subprime Lending

Adam Rust
Research Director
Community Reinvestment Association of North Carolina
Durham, NC 27701
adam@cra-nc.org

HMDA data manifests the ambition that prompted and helped to pass the Community Reinvestment Act. One person said it this way: “We have learned from 30 years of CRA policy that what is measured gets done.”³

The data is important, but even in saying that, it is important to acknowledge that there is room for improvement. Mortgage lending has changed a great deal in the last decade, but HMDA data has introduced only one substantial improvement in that time. I am encouraged by the revisions to HMDA data collection that are expected in the current version of the Dodd-Frank Wall Street Reform and Consumer Protection Act.

I would like to submit a paper that was published in connection with the conference “Revisiting the CRA: Perspectives on the Future of the Community Reinvestment Act,” with my written comments.

I am very pleased to see that The Dodd-Frank bill will satisfy many of those concerns. Here is a list of some of those ideas:

- Loan-to-value
- Credit score
- Length of introductory-rate period
- New rate spread formula
- Origination channel
- Term to maturity
- Borrower age
- Prepayment penalty
- Presence of a negative amortization feature
- Points and fees

By making these data explicit requirements for collection, Congress has given HMDA users more ability to answer the questions that are posed to them by the community members that they represent. Implementing data rules for the last five features are clear cut. There is no need to clarify the age of a borrower, for instance. With the first five, however, the construction of

the data remains open for interpretation. The details of how the indicators are constructed will determine their value for advocates. I will include some comments on how Congress' data plans could be best implemented.

These are helpful. More could be done.

Banks continue to tell me that HMDA data does not offer a legitimate portrayal of their lending. Here is an excerpt from a letter sent to me by Ameriquest (ACC Capital Holdings) in May 2005, after the new rate spread information was included in new reporting rules.

“The information generated by HMDA’s reporting requirements is important in evaluating mortgage lending patterns nationwide. We note, however, that the HMDA data also has serious shortcomings, as it does not include key financial information that is essential to analyzing the fairness of the rate associated with a given loan. This important information might include credit score, property type, down payment, ash-out information, property value, debt-to-income ratio, loan-to-value ratio and assets. In short, such exclusions mean that HMDA data alone cannot be used to draw any conclusions about why a loan was refused or why it was priced at a particular rate^b.”

This letter is unusual for the specificity of its critique. It lists discrete data points that are missing from HMDA. Other lenders routinely downplay the value of HMDA, but they are usually dismissive in a general matter. Ameriquest went the extra mile.

More broadly, the important point is to hew to principles that can make the data more applicable. First, HMDA data should be redesigned to tell more about the quality of lending. Second, efforts should be made to make the data easier to use. Third, expand the coverage for other segments of the mortgage market. Last, add more detail to small business data.

Advocates and the communities that they represent would benefit if the Federal Reserve made additional fixes. It would reflect the new dynamics that are common with mortgage markets. Subprime lending has made capital far more available than it was thirty years ago, but it has also brought significant risks for the very households. Data can make a difference. It might have helped more to thwart the recent crisis. Going forward, these changes can help to alert the public about future risks.

Uses for HMDA Data

Community groups are taken much more seriously when they can complement their direct dialogue with a financial institution with verifiable statistics. It is one thing to say “you don’t make loans in our neighborhood,” and quite another to say “you only made two prime rate loans in our census tract in the last three years, and both were for investors.” The former is merely an opinion, while the latter is given authority and nuance by its use of factual data.

I use HMDA data in many ways. We analyze the availability of capital by reviewing the share of loans that denied. We judge the quality of capital by the frequency of loans that exceed the rate spread threshold.

The common purpose of our questions, posed since the CRA was passed, is ‘how are banks and thrifts meeting the credit needs of their local communities’? That is still a frequent impetus for the use of this data. I’ve answered that question not just for local community groups, but also for Mayors and County Commissioners. We routinely present our conclusions on the story told by HMDA data to the banks that have submitted the data.

Still, the data has other uses. I have used the data to construct an index for socially responsible investors. A national church group has asked us to use HMDA data in a project that interacts with data from the FDIC, the National Information Center, and other government data sources.

Of those “other uses,” the most common questions ask us to distinguish between “good” loans and “bad” loans. The “old HMDA” is not designed for that question. It assumed that most mortgage loans were the same. More loans were good, and fewer loans were bad. That viewpoint, in the light of what we know now in 2010, is optimistic at best. More often than not, the new focus is qualitative. The traditional question was ‘how much credit?’ That has been replaced by a new concern, namely ‘what type of credit?’ How can we create data that will communicate the quality of lending as clearly as possible?

Many researchers use HMDA data to develop arguments about the safety and soundness of lending. Unfortunately, the most authoritative reports depended upon analyses that used private data to complement HMDA. One of HMDA’s appeals is its cost. Data from First American, Core Logic, or LPS all bear costs that exclude most users.

How Congress’ Expectations Should be Implemented

Loan-to-Value: Include a reference for loans with a cash-out feature. Cash-out would have to be paired with a separate LTV figure. It might be stated in terms of cash-out amount (in thousands of dollars) or as a percentage of the loan amount. Assuming that loan-to-value will be expressed as a ratio, it probably makes sense to do the same with a cash-out value.

In a similar vein, the loan-to-value indicator should also include some indicator for the presence of a second loan in a home purchase. Many borrowers used a second loan to avoid paying mortgage insurance. This was a risky practice and one that HMDA could help to spotlight.

Credit Score: Create a categorical credit score variable. Congress intends to let the Bureau determine the scope of the credit score. The ideal data point here achieves a compromise between the needs of users to control for credit quality and the right of consumers to have some degree of privacy. A credit score data point should be a categorical variable with perhaps

four credit bands: two for different grades of subprime credit, one for median credit, and one for prime credit. This reduces the noise of identify a consumer by a falsely exact score, and it creates a relative “cloak” of privacy for consumers. Vantage Score already provides this kind of metric. It collapses credit scores into a range of six letter grades.

Length of Introductory Period: There is more than one relevant descriptor for the terms of an adjustable-rate loan. Some loans are marketed with interest rates that reset in just 30 days. The length of introductory period will communicate the nominal status of a loan. It is either adjustable or fixed. However, there are more relevant details. Adjustable-rate loans vary by their subsequent reset periods as well as by their maximum interest rates. Those should be accounted for as well.

A new rate spread formula: Currently, rate spreads are evaluated against the initial price of a loan. In years with higher yield curves, this design allows most adjustable-rate loans to escape the rate spread threshold. If an adjustable-rate indicator is developed, it prompts a new need for a means of assessing the cost of that loan. I would argue that just as the “comparably termed Treasury” serves to judge the price of an existing loan, so might the cost of a Treasury Inflation Protected Security (TIPS). This month, the 10-year TIPS Auction yielded 1.43 percent and the 5-year TIPS yielded 0.55 percent. By comparison, a 10-year fixed note yielded 3.55 percent. The same 300 basis point standard, if re-applied to a TIPs with a similar term, would give analysts a better sense of the risk-adjusted price of new adjustable-rate mortgage loans.

Origination Channel: The proposed CFPB changes will require loans to indicate if a loan comes through a mortgage broker or traditional retail channel. The broker flag is particularly relevant. Mortgage brokers actively participated in subprime lending. More than eighty percent of subprime loans originated in California in 2005 and 2006 were channeled through a mortgage brokers^c. It will be valuable to have new data about the channel, but there are more ways that this indicator could be redesigned to add to ease of use for analysts.

The data should reveal the corporate parent for all loan originations. More often than not, analysts want to assess the lending of a corporation, rather than just one channel. This would add to accuracy, as well as to ease of use. Wells Fargo, for instance, issues mortgages through more than 60 channels. As it is currently constructed, it can be hard to be sure that the right LARs have been identified. This will remain important into the future. Mergers and FDIC-facilitated sales will make this a common problem in the near future. It is relatively easy to append Wachovia loans from 2008 into Wells Fargo.

The data should also identify when loans are provided through a relationship with a home builder. Many lenders used in-house financing as an incentive for the sale of newly constructed homes. Beazer, Centex, Standard Pacific, Ryland, and Pulte all have mortgage subsidiaries. The problem with in-house lenders is that it can create too much “origination risk” in a neighborhood. When all of the loans in a subdivision have a similar vintage and bear a narrow

window of resets, then the community is vulnerable to a sudden crop of foreclosures. If the loans have subprime features, the risk is even greater.

The origination channel should also indicate how that loan fits into a CRA assessment. In 2008, I participated in a paper that found that CRA-regulated banks issued high-cost loans much more frequently when they operated outside of the areas in the branch network. This was a two-step process. First, we had to identify which channels were obligated for CRA. Then we had to filter those applicable loans by their MSA.

Additional Changes that Would Help

There are additional sectors of mortgage markets that deserve attention.

Manufactured Housing: In North Carolina, manufactured housing is a vital portion of our housing stock. It serves one in six households across the state, and most likely a far higher share of low-and-moderate income households. HMDA is unclear on how it differentiates between homes financed as personal property and those financed as real property. I have sought clarification from the Federal Reserve and I have received conflicting answers. Some say that only those classified as real property are in the data. Others tell me that all properties used as a habitation are reported. In the future, the data should make this distinction clear.

Net Tangible Benefit for Refinance Loans: The Federal Reserve should construct a net tangible benefit formula for refinance loans. The CFPB will most likely require this as a new consumer protection. If included in HMDA data, this indicator would enhance the fidelity between HMDA and the concerns of regulators.

Debt-to-Income: This is an important tool for evaluating how lending will influence the stability of property values in neighborhoods. When borrowers enter foreclosure, the distressed sale establishes a low “comparable” for surrounding homes^d. It also impacts the ongoing tax base of the local municipality^e. A debt-to-income indicator would help analysts. It would be most relevant for predicting loan performance. The indicator would ideally be broken out to indicate both the DTI that is exclusively a product of the mortgage, and then a second figure that captures the total percentage of income that is devoted to servicing all debt. This data would inform loan modification research as well. Borrowers exiting HAMP with a permanent modification have a 31 percent front-end debt-to-income ratio, but on average, they still are burdened by too much debt. The average permanent loan modification recipient had a back-end debt-to-income ratio of 61.3 percent^f.

Stated-Income or No-Documentation Loans: Data should flag loans where income was not verified. Currently, many loans lack income data, but that is not a satisfactory means for analysts to gauge the underwriting of loans. This was one of the riskiest innovations in the recent subprime boom. While it is rarely applied to underwriting today, it may come back in the future. By

identifying loans with this feature, researchers could play a role in sounding alerts about the safety of new loans.

Loan Purpose: The existing loan purpose indicator is not capable of expressing the full spectrum of loan purposes. It does not identify home equity loans. This purpose should be added.

I believe that users would find a benefit for an expansion of the home purchase category. It would be an enhancement to distinguish between home purchase loans used to buy a newly constructed home and for those used to buy an existing home. One of the motivations for the CRA was to make sure that banks were providing capital to all neighborhoods. A lack of lending for home sales in existing neighborhoods undergirded the contentions of advocates at the time. It would allow advocates to better capture the rate of reinvestment in older neighborhoods. There would also be demand for more information about new construction. This is a strong driver of job creation.

Add geographic coordinates. In the future, more analysis will combine the power of databases with the capacity of geographic information systems. GIS can make some use of HMDA data. It can aggregate loan activity by a geographic area, such as census tract or MSA, in order to make some analysis of loan data. Census tract is the smallest level of geography. While this is a fairly narrow area, census tracts are not concurrent with neighborhoods. Most often, a census tract contains about two thousand people. Some neighborhoods are within one census tract, while others are spread out across several.

The solution would be to append data with x and y coordinates that would allow the creation of vector data. With that capacity, analysts could identify instances when specific neighborhoods have been victimized by poor lending practices.

Small Business Data: Very little research comes from the FFIEC's publication of CRA data. While there is not widespread interest in farm lending, many people would like to have a better means for understanding the provision of small business loans in their community.

The most significant change that could be made would be to re-organize small business lending onto a loan-by-loan data format. The next table includes a brief summary of potential enhancements to the small business lending database. The relevant category suggests the motives for that element. A lack of uniformity hints at some challenges for that design.

CRA small business lending data should follow the spirit of HMDA’s mortgage design. HMDA data is offered both at the loan level and in aggregated reports. Given the choice, almost all users prefer loan level information. There is no relief to be had by paying for alternative data.

<i>Small Business Variable</i>	<i>Outcomes</i>	<i>Relevant?</i>	<i>Uniform?</i>
LOAN			
Loan purpose	capital expenditure, inventory, working capital	Heterogeneity	No
Loan decision	Originate, approve, deny, incomplete	Fair lending	Yes
Loan term	Categorical term length, or line of credit	Heterogeneity	Yes
Collateralization	Equity, real property, inventory, personal, other, none	Heterogeneity	Yes
Loan amount	Specify amount	Clarity	Yes
BORROWER			
Business type	Three-digit NAICS classification	Heterogeneity	Yes
Debt to equity	Liabilities/equity	Ability to repay	Yes
Working capital	Current assets/current liabilities	Ability to repay	Yes
Owner designation	Identify minority or female-owned business	Fair lending	Yes
Revenue	Maintain in new database	Ability to repay	Yes
Franchisee	Yes/no	Management	Yes
Firm size	Categorical indicator of number of employees	Job creation	Yes
Firm experience	Categorical indicator of firm tenure	Job creation	Yes
Job creation	No, or quantity of jobs	Job creation	Yes

The Small Business Administration has some data. Unlike mortgages, there are few private suppliers of this information. In the absence of that data, analysts have little to go on to assess small business volume.

Some Data is Not Relevant

Some data points in the existing HMDA LAR have outlived their usefulness. I have never found a reason to use the “number of owner occupied units” variable or the “number of 1 to 4 Family Units” variable. In five years, I have only used total population in tract once.

In the Near Future

Adding these credit capacity, underwriting, and loan feature data points would bring immediate help to community groups. The next big issue for these groups will be the new add-on fees that are appended to the securitization of home mortgages. The GSEs and FHA have implemented new policies that could have the unintended consequence of adding to the disinvestment in low-income neighborhoods. The Loan Level Pricing Adjustment (LLPA) and the Adverse Markets Delivery Charge (ADMC) add costs for lenders seeking to resell loans. Inevitably, these costs will be passed on to consumers.

While the intent of the LLPA and ADMC is largely to maintain the solvency of the GSEs, it could have an unintended consequence of increasing the cost of borrowing for low-income communities. The LLPA and the ADMC trigger additional origination fees for loans with low

credit scores, high debt-to-incomes, and high loan-to-value measures. These are all factors that will impact the availability of credit.

The LLPA holds the possibility of limiting the supply of credit in traditionally underserved areas. Its design has no explicit bias, but it is likely that the burden of its impact will be most severe in low-income neighborhoods and to minority and low-income borrowers. This is driven by its use of credit score and down payment as inputs for pricing.

Research has documented that there are significant discrepancies in the median credit scores of non-Hispanic white borrowers compared to those of African-American and Hispanic households^g. The impacts even extend to entire areas. The South, as a whole, has a mean credit score that is 21 points lower than the rest of the country. Counties with higher rates of minority residents are also more likely to have a lower average credit score^h. The new costs in these policies are going to fall on the backs of minority and low-income communities. The LLPA will impose minimal costs on borrowers with credit scores about 740. For borrowers with scores that are lower, the incremental costs are steep.

By adding fees for loans with high balance LTVs, it creates an additional hurdle for borrowers with fewer assets. Many low-income and minority borrowers utilize the FHA program to overcome requirements for high down paymentsⁱ.

It is a classic CRA issue. Unfortunately, HMDA data, as currently constructed, is almost useless for the task of gauging this policy. Only three of the 13 factors in the LLPA are found within the existing HMDA LAR.

Indicator	In HMDA?	In LLPA?	Basis Points	Factor
ARM	No	Yes	Up to 25	LTV
Balloon Mortgage	No	Yes	12.5	LTV
Investment Property	Yes	Yes	175 to 375	LTV, Balloon
Multiple Unit Property	Yes	Yes	Up to 100	LTV
Manufactured Home	Yes	Yes	50	LTV
40 Year Term	No	Yes	125	LTV
Interest-Only	No	Yes	25 to 100	LTV, Balloon
Condominium or Co-Operative	No	Yes	Up to 75	LTV, Term, Balloon
Cash-Out Refinance	No	Yes	0 to 300	Credit Score
High Balance (LTV)	No	Yes	75	ARM, Cash-Out Refi
Credit Score	No	Yes	Up to 300	LTV
Subordinate Financing	Yes	Yes	25 to 75	Credit Score, I/O, LTV, CLTV
My Community Mortgage	No	Yes	75 to 125	Subordinate Financing, I/O, Term, ARM

Source: Fannie Mae

The reforms that are needed for HMDA data would answer the concerns of someone trying to model the impact of the LLPA and the AMDC.

The implicit assumption behind the LLPA is that these additional factors are likely to increase risk. These are safety and soundness factors. The policy governs loans delivered to the agencies. However, it is hard to imagine that the new policies will not carry over into origination fees

paid by retail consumers. The consequence of this new dynamic will certainly fall heavily on low and moderate income constituencies.

Conclusion

HMDA data could help to identify problems in mortgage markets.

Its redesign will also help to fulfill the intentions of the Home Mortgage Disclosure Act and the Community Reinvestment Act. We want data that does not just observe the letter of the law, but that goes the extra mile to achieve a standard that honors the spirit of that legislation.

The larger lesson that can be drawn from this example should be clear. In an era of subprime lending, HMDA data no longer works. The new language in the proposed CFPB could go a long way toward reinvigorating the role of HMDA. The data still asks for more fixes, though. It can be a great tool.

HMDA data needs to be flexible. It is a natural challenge for regulators to stay even with the pace of innovation. Lenders adopted risk-based pricing, loosened their credit standards and developed new loan products in the last fifteen years. With one exception (the rate spread indicator), HMDA remained the same. It is time to update.

References

- ^a Ellen Seidman, New America Foundation. Testimony before the Committee on Financial Services, U.S. House of Representatives, 2008, available at <http://www.newamerica.net/files/CRA%20Testimony%202-13-08.pdf>.
- ^b ACC Capital Holdings. Letter to Adam Rust. May 5, 2005.
- ^c Leonard, Paul. 2008. *The Subprime Mortgage Crisis: Context, Causes, Solutions*. Center for Responsible Lending. http://www.frbsf.org/community/issues/assets/preservation/resources/crl_presentation.pdf
- ^d Immergluck, Dan and Geoff Smith. 2006. *The External Costs of Foreclosure: The Impact of Single-Family Mortgage Foreclosures on Property Values*. *Housing Policy Debate*. 17(1)57-79.
- ^e Apgar, William, Mark Duda, Rochelle Gorey. 2005. *The Municipal Costs of Foreclosures: A Chicago Case Study*. A Report Prepared for the Homeownership Preservation Foundation.
- ^f Making Home Affordable Program. *Servicer Performance Report*. March 2010.
- ^g Board of Governors of the Federal Reserve System. 2007. "Report to the Congress on Credit Scoring and its Effects on the Availability and Affordability of Credit." <http://www.federalreserve.gov/boarddocs/rptcongress/creditscore/creditscore.pdf>
- ^h Fellows, Matt. 2006. "Credit Scores, Reports, and Getting Ahead in America." Brookings Institution. http://www.brookings.edu/~media/Files/rc/reports/2006/05childrenfamilies_fellows/20060501_creditscores.pdf
- ⁱ Monroe, Albert. 2001. *How the Federal Housing Administration Affects Homeownership*. Joint Center for Housing Studies. Cambridge, Massachusetts.