HMDA Sampling Procedures

The following HMDA sampling procedures should be applied when reviewing HMDA-LAR data for accuracy:

1. **Identify and select the LAR to be reviewed**.

For each HMDA reporter, review both current year and previously submitted HMDA data since the last examination. Every examination conducted after April 30th each year should include a review of the current year's HMDA data. Examinations conducted before April 30th should include a review of the current year's data to the extent that the institution has already entered data from the current year on the LAR. The data from a single year's LAR is the universe from which a sample is taken.

2. **Determine the total number of files to be sampled from Column G on the HMDA Sampling Schedule based upon the universe size.**

For banks at which HMDA data is not relied upon in conducting the fair lending or CRA examination, the product module and examination matrix may indicate a Level II review, involving sampling as appropriate. In these instances, examiners should choose a judgmental sample that is sufficiently large to ensure confidence in the overall accuracy of the data.

3. **Select the total random sample.**

A. <u>From an Automated Download</u>:
   The most important thing to remember is the sample must be <u>randomly</u> selected from the universe. There are a variety of tools an examiner may use to electronically select a random sample of data, including a feature in Excel. The following instructions will assist you in working with Excel:

   1. Generate a random order to the universe of records from which the sample will be selected. This generation should be made by utilizing Excel's "Random Number Generation" tool according to the following steps:
      a. Select the following from the Excel menu:
         - Tools
         - Add-Ins
         - Analysis Tool Pak (Check box and click "OK")
         - Tools (Again)
         - Data Analysis
         - Random Number Generation (Highlight and click "OK")
      b. Respond to the items on the Random Number Generation screen as follows:
         - Number of Variables (Leave blank)
         - Number of Random Numbers (Leave blank)
         - Distribution (Select "Uniform" from list)

- Parameters (Leave default as is - it is set at 0 and 1)
- Random Seed (Leave blank)
- Output Options (Click on "Output Range" circle and then click on the small box to the right for "Output Range")

   c. A small screen will appear titled "Random Number Generation" but you will not enter any information directly onto that screen. Rather, select the range (output location) for the random numbers by highlighting the column on the spreadsheet where you want the random numbers to go. (Use the "Shift" key and the down arrow to highlight the column.) Hint: Designating a column at the end of the spreadsheet may be easiest.

   d. Click on the small box in the "Random Number Generation" screen (or press Enter).

   e. Click on "OK".

   f. The random numbers are automatically assigned and placed into the designated statistical column.

   g. Sort the file in ascending order by the random number by highlighting all of the data, selecting "Data", then selecting "Sort", then identifying the column (containing the random numbers) by which you will sort, then selecting "Ascending", and finally, selecting "OK".

2. Once the loans are placed in a random order, the sample needed for the HMDA verification is simply taken from the top of the list. This information should then be saved as a supporting workpaper.

B. From Hard-Copy LAR:

   As with electronic data, a sample of files selected from a manual HMDA report must be <u>randomly</u> selected from the universe.

1. Divide the size of the ''universe'' by the size of the sample to determine the ''interval.'' If necessary, round interval down to reach a whole number.
2. Randomly pick a number between zero and the interval.
3. Starting with the beginning of the universe, count the items until the random number is reached. The item that corresponds with the random number will be the first item selected for the sample.
4. Starting with the next consecutive item as number 1, count until the number corresponding with the interval is reached and select that item for the sample.
5. Repeat Step 4. above throughout the entire universe, until the appropriate sample size is reached.

4. Review the initial number of files (Column B in the HMDA Sampling Schedule at the end of this section) utilizing current FRB HMDA data review procedures.

5. The examiner may stop the HMDA sampling process after review of the initial number of files is completed IF the results indicate that a very small number of files had errors in key fields[1]. Using the HMDA Sampling Schedule, this number can be determined by referencing Column C titled "Maximum Number of Files with Errors to Stop Sampling."

---

[1]**Key fields** are defined as: loan type; loan purpose; property type; owner occupancy; loan amount; action taken type; request for preapproval; application date and action date; MSA; state; county; census tract; ethnicty, race, and sex of the applicant and co-applicant; income; type of purchaser; rate spread; HOEPA status; and lien status.

For example, if a HMDA universe contains 150 files, a total random sample of 56 files should be taken. The examiner may initially begin file review on 29 files. If, upon completing review of the initial 29 files, the examiner finds no more than one file with any error or errors in key fields, the examiner may end the sampling process for that HMDA reporter for that universe. The examiner may then reach a statistically reliable conclusion that the findings are indicative of the universe and resubmission is not necessary.

6. **The examiner must complete a review of the total random sample of files if a larger number of errors in key fields are found during the initial file review.**

The need for this additional file review can be determined by using Column D titled "Number of Files with Errors - Additional File Review Required" on the HMDA Sampling Schedule. If the number of files with errors in key fields from the initial review falls within the number reflected in this column, the examiner must review the additional files to complete the total random sample.

For example, if a HMDA universe contains 150 files, a total random sample of 56 files should be taken. The examiner may initially begin file review on 29 files. If, upon completing review of the initial 29 files, the examiner finds four files with an error or errors in key fields, the examiner should then review 27 additional files, for a total sample size of 56 files. After completing review of the additional 27 files, the examiner should determine the total number of key field errors and apply the current Board HMDA resubmission standards to the total sample.

7. **If an examiner determines that a large number of files reviewed in the initial file review have an error or errors in key fields, the examiner may stop HMDA data verification after the initial file review is completed and should apply the current Board HMDA resubmission standards.**

This maximum number can be determined by referencing Column E titled "Number of Files with Errors to Stop Sampling - Apply Resubmission Standards on the HMDA Sampling Schedule." For example, if a HMDA universe contains 150 files, a total random sample of 56 files should be taken. The examiner may initially begin file review on 29 files. If, upon completing review of the initial 29 files, the examiner finds six (or more) files with an error or errors in key fields, the examiner should stop file review. Sufficient statistical evidence has been obtained to conclude that a larger sample would have an unacceptable number of errors, thus requiring resubmission. At this point, the examiner should apply the current Board HMDA resubmission standards to the total sample.

HMDA Sampling Schedule

A	B	C	D	E	F	G
		INITIAL FILE REVIEW				
HMDA UNIVERSE	Initial File Review	Maximum Number of Files with Errors* - Stop Sampling	Number of Files with Errors* - Additional File Review Required (go to column F)	Minimum Number of Files With Errors* - Stop Sampling & Apply Resubmission Standards	ADDITIONAL FILE REVIEW	TOTAL RANDOM SAMPLE
1-12	Review all					
12-20	12	0	1	2	Review all	All
21-30	13	0	1	2	Review all	All
31-50	15	0	1-2	3	13	28
51-70	17	0	1-2	3	12	29
71-90	18	0	1-3	4	20	38
91-110	28	1	2-3	4	11	39
111-130	29	1	2-4	5	18	47
131-140	29	1	2-4	5	20	49
141-170	29	1	2-5	6	27	56
171-190	30	1	2-5	6	27	57
191-270	30	1	2-5	6	29	59
271-380	30	1	2-6	7	38	68
381-750	31	1	2-6	7	38	69
751-1100	31	1	2-7	8	48	79
1101-	32	1	2-7	8	47	79

*   Files with one or more errors in Key Fields (Key fields identified by the Federal Reserve Board are: loan type; loan purpose; property type; owner occupancy; loan amount; action taken type; request for preapproval; application date and action date; MSA; state; county; census tract; ethnicty, race, and sex of the applicant and co-applicant; income; type of purchaser; rate spread; HOEPA status; and lien status).