"NATIONAL SURVEY OF MORTGAGE ORIGINATIONS"

OMB NUMBER 2590-0012

SUPPORTING STATEMENT

The Federal Housing Finance Agency (FHFA or the Agency) is seeking approval for a three-year extension of the Paperwork Reduction Act (PRA) clearance for the National Survey of Mortgage Originations (NSMO). OMB has assigned the NSMO control number 2590-0012, which is due to expire on December 31, 2016. This information collection was called the "National Survey of Mortgage Borrowers" when it was last submitted for PRA clearance in 2013. FHFA changed the name of the survey to reflect more accurately the focus of the questions and to distinguish it more clearly from the American Survey of Mortgage Borrowers (OMB control no. 2590-0015).

The NSMO is a recurring quarterly survey of individuals who have recently obtained a loan secured by a first mortgage on single-family residential property. The survey questionnaire is sent to a representative sample of approximately 6,000 recent mortgage borrowers each calendar quarter and typically consists of between 90 and 95 multiple choice and short answer questions designed to obtain information about borrowers' experiences in choosing and in taking out a mortgage.[1] The NSMO is sponsored by FHFA and is one component of the National Mortgage Database Project, an ongoing joint effort of FHFA and the Consumer Financial Protection Bureau (CFPB).

Many of the topics addressed in this Supporting Statement—in particular those relating to the statistical methods addressed in Part B—are discussed in greater detail in the November 14, 2016 draft of NSMO Technical Report 2.1, prepared by FHFA's National Mortgage Database Project staff, which is included as Attachment 2.

A. JUSTIFICATION

1. Circumstances necessitating the collection of information

Section 1324 of the Federal Housing Enterprises Financial Safety and Soundness Act of 1992 (Safety and Soundness Act) requires that FHFA prepare annually a detailed report on the residential mortgage market activities of two of its regulated entities—the Federal National Mortgage Association (Fannie Mae) and the Federal Home Loan Mortgage Corporation (Freddie Mac) (collectively, "the Enterprises")—and to submit that annual report to the Committee on Banking, Housing, and Urban Affairs of the Senate and the Committee on Financial Services of the

---

[1] A copy of the current version of the NSMO questionnaire (mailed in November 2016) is included as Attachment 1 to this Supporting Statement. Copies of the questionnaire in both English and Spanish can also be accessed online at: http://www.fhfa.gov/Homeownersbuyer/Pages/National-Survey-of-Mortgage-Originations.aspx.

House of Representatives.[2] At a minimum, the report must: (1) address the extent to which the Enterprises are fulfilling their statutory duties with respect to the residential mortgage markets, including their duty to serve underserved markets; (2) aggregate and analyze relevant data on income to assess the compliance of each Enterprise with statutory housing goals established under section 1331 of the Safety and Soundness Act;[3] (3) aggregate and analyze data on income, race, and gender by census tract and other relevant classifications, and compare such data with larger demographic, housing, and economic trends; (4) identify the extent to which each Enterprise is involved in mortgage purchases and secondary market activities involving subprime and nontraditional loans; (5) compare the characteristics of subprime and nontraditional loans purchased and securitized by each Enterprise to other loans purchased and securitized by each Enterprise; and (6) compare the characteristics of high-cost loans purchased and securitized, but not held in portfolio, by each Enterprise to such securitized loans that are retained in portfolio or repurchased by the Enterprise, including such characteristics as the purchase price of the property securing the mortgage, the loan-to-value ratio of the mortgage, the terms of the mortgage, the creditworthiness of the borrower, and any other relevant data, as determined by the Director of FHFA.[4]

Section 1324 further requires that FHFA conduct a monthly survey to collect data needed to adequately analyze the matters that must be addressed in the annual report.[5] In particular, the survey must collect information on the characteristics of individual prime and subprime mortgages and the creditworthiness and other characteristics of the borrowers on those mortgages.[6] It may also address such other matters as the Director of FHFA deems to be appropriate.[7] The statute requires that FHFA compile a database of timely and otherwise unavailable residential mortgage market information obtained from the monthly survey and to make that information available to the public.[8]

As a means of fulfilling these and other statutory requirements, as well as to support policymaking and research regarding the residential mortgage markets, FHFA and CFPB jointly established the National Mortgage Database Project in 2012. The project is designed to provide comprehensive information about the U.S. mortgage market based on a five percent sample of residential mortgages. The project has three primary components: (1) the National Mortgage Database; (2) the quarterly NSMO; and (3) the annual American Survey of Mortgage Borrowers (ASMB).[9] When fully complete, the National Mortgage Database will be a de-identified loan-level database of closed-end first-lien residential mortgage loans that is representative of the market as a whole, contains detailed loan-level information on the terms and performance of the mortgages and the

---

[2] *See* 12 U.S.C. 4544(a). Congress added the requirements of section 1324 to the Safety and Soundness Act in 2008. *See* Housing and Economic Recovery Act of 2008, Pub. L. No. 110-289, sec. 1125, 122 Stat. 2654, 2693-95 (2008). A copy of section 1324, as amended, is included as Attachment 3.
[3] 12 U.S.C. 4561.
[4] *See* 12 U.S.C. 4544(b).
[5] *See* 12 U.S.C. 4544(c)(1).
[6] *See* 12 U.S.C. 4544(c)(2)(A), (B).
[7] *See* 12 U.S.C. 4544(c)(2)(C).
[8] *See* 12 U.S.C. 4544(c)(3).
[9] While the NSMO solicits information about the experiences of borrowers who have recently obtained a mortgage, the ASMB solicits information on borrowers' experience with maintaining their existing mortgages. OMB has cleared the ASMB under the PRA and assigned it control no. 2590-0015, which expires on July 31, 2019.

characteristics of the associated borrowers and properties, is continually updated, has an historical component dating back to 1998, and provides a sampling frame for surveys to collect additional information.

The core data in the National Mortgage Database are drawn from a random 1-in-20 sample of all closed-end first-lien mortgage loans outstanding at any time between January 1998 and the present from the files of Experian, one of the three nationwide credit reporting agencies. The National Mortgage Database currently contains data on approximately 11.6 million mortgage loans. Between 80,000 and 100,000 mortgages, drawn from a random 1-in-20 sample of loans newly reported to Experian, are added each quarter. Additional information on the mortgages in the datasets is drawn from other existing sources, including, but not limited to the Home Mortgage Disclosure Act (HMDA) data released by the Federal Financial Institutions Examination Council (FFIEC), property valuation models, transactional data maintained by local governments, and administrative data files maintained by the Enterprises and by federal agencies. Mortgages are followed in the National Mortgage Database until they terminate through prepayment (including refinancing), foreclosure, or maturity.

The NSMO was developed to complement the National Mortgage Database by providing critical and timely information—not available from existing sources—on the range of nontraditional and subprime mortgage products being offered, the methods by which these mortgages are being marketed, and the characteristics of borrowers for these types of loans. In particular, the survey questionnaire is designed to elicit directly from mortgage borrowers information on the characteristics of borrowers and on their experiences in finding and obtaining a mortgage loan, including: their mortgage shopping behavior; their mortgage closing experiences; their expectations regarding house price appreciation; and critical financial and other life events effecting their households, such as unemployment, large medical expenses, or divorce. The survey questions do not focus on the terms of the borrowers' mortgage loans because these fields are available in the Experian data. However, the NSMO collects a limited amount of information on each respondent's mortgage to verify that the Experian records and survey responses pertain to the same mortgage.

Each wave of the NSMO is sent to the primary borrowers on about 6,000 mortgage loans, which are drawn from a simple random sample of the 80,000 to 100,000 newly originated mortgage loans that are added to the National Mortgage Database from the Experian files each quarter (at present, this represents an approximately 1-in-15 sample of loans added to the National Mortgage Database and an approximately 1-in-300 sample of all mortgage loan originations).[10] Wave 1 of the NSMO was mailed out in April 2014, and a new wave of the survey has been conducted each quarter since—to date, eleven quarterly waves of the survey have been completed. Initial analysis of the data from Waves 1-3 of the survey resulted in the addition and removal of some questions and the

---

[10] The first two waves of the survey varied from the ultimate target number of 6,000 survey recipients. Wave 1 of the survey, mailed in April 2014, included a sample of 15,000 mortgages. This was a catch-up period to cover cases originated in 2013 and newly reported to Experian in the archives for June, September, and December 2013. Wave 2, mailed in June 2014, included a sample of 3,000 mortgages, covering mortgage loans that were originated in 2013 and newly-reported to Experian between January and March 2014.

revision of several other questions. These revisions were incorporated into the survey questionnaire for Wave 7 of the NSMO, which was conducted in the summer of 2015.

The NSMO grew out of a series of three pilot surveys that were funded and carried out by Freddie Mac between 2010 and 2012. The NSMO questionnaire draws on the questionnaires piloted by Freddie Mac and leverages the input of an advisory group of industry experts from government, non-profits, advocates, trade groups, and academia.

By interagency agreement between FHFA and CFPB, FHFA has led the production and execution of the NSMO.[11] By contract with FHFA, the conduct of the NSMO is administered through Experian, which has subcontracted the survey administration through a competitive process to Westat, a nationally-recognized survey vendor.[12] Westat also carries out the pre-testing of the survey materials.

2. Use of data

FHFA views the National Mortgage Database Project as a whole, including the NSMO, as the monthly "survey" that is required by section 1324 of the Safety and Soundness Act. Core inputs to the National Mortgage Database, such as a regular refresh of the Experian data, occur monthly, though NSMO itself does not. In combination with the other information in the National Mortgage Database, the information obtained through the NSMO is used to prepare the report to Congress on the mortgage market activities of Fannie Mae and Freddie Mac that FHFA is required to submit under section 1324, as well as for research and analysis by FHFA and CFPB in support of their regulatory and supervisory responsibilities related to the residential mortgage markets. The NSMO is especially critical in ensuring that the National Mortgage Database contains uniquely comprehensive information on the range of nontraditional and subprime mortgage products being offered, the methods by which these mortgages are being marketed and the characteristics—and particularly the creditworthiness—of borrowers for these types of loans. In the future, the information may be used to provide a resource for research and analysis by other federal agencies and by academics and other interested parties outside of the government.

FHFA will use information collected through cognitive pre-testing of survey materials to assist in drafting and modifying the survey questions and instructions, as well as the related communications, to read in the way that will be most readily understood by the survey respondents and that will be most likely to elicit usable responses. Such information will also be used to help the Agency decide on how best to organize and format the survey questionnaires.

---

[11] An interagency agreement between FHFA and CFPB was signed on September 12, 2012 where the costs of the survey and the development of the National Mortgage Database are to be shared equally between the two agencies.
[12] The Fair Credit Reporting Act, 15 U.S.C. 1681 *et seq.*, requires that the survey process, because it utilizes borrower names and addresses drawn from credit reporting agency records, must be administered through Experian in order to maintain consumer privacy.

3. Use of information technology

The NSMO uses machine-readable paper questionnaires and also gives recipients the option of completing the survey online in either English or Spanish. Completed paper questionnaires are scanned and the responses are automatically uploaded into the electronic National Mortgage Database.

4. Efforts to identify duplication

As explained above, the vast majority of data included in the National Mortgage Database is drawn from existing sources—primarily, the consumer credit database maintained by Experian; the HMDA data released by FFIEC; and administrative data in the possession of FHFA, its regulated entities, and other federal agencies. As described under Item #1, the NSMO questionnaire is designed to obtain critical and timely information that is not available from existing sources. The survey obtains this information directly from borrowers, who are likely to be the most reliable and accessible—and, in some cases, the only—source for this information.

5. Impact on small entities

The information collection will not have a significant economic impact on a substantial number of small entities. The survey recipients are individuals only.

6. Consequences of less frequent collection and obstacles to burden reduction

Section 1324 of the Safety and Soundness Act requires that FHFA undertake a survey of mortgage markets on a monthly basis.[13] While the performance data on existing mortgages in the National Mortgage Database is pulled from the Experian database on a monthly basis, newly originated mortgages are added to the National Mortgage Database on a quarterly basis. The NSMO questionnaires are sent to a random sample of borrowers that originated their mortgages in the year and quarter that corresponds to the quarterly draws of newly originated mortgages from the Experian database. One important purpose of the survey is to monitor loan origination trends. While monthly housing surveys would provide the optimal feedback regarding these trends, FHFA believes that quarterly surveys are sufficient.

7. Circumstances requiring special information collection

There are no special circumstances that require FHFA to conduct the information collection in a manner inconsistent with the guidelines provided in this Item 7.

8. Solicitation of comments on information collection

In accordance with the requirements of 5 CFR 1320.8(d), FHFA published a request for public comments regarding this information collection in the *Federal Register* on September 13, 2016.[14]

---

[13] *See* 12 U.S.C. 4544(c).
[14] *See* 81 FR 62889 (Sept. 13, 2016). A copy of the 60-day Notice is included as Attachment 4.

The 60-day comment period closed on November 14, 2016. FHFA received two comment letters, one of which recommended revisions to the content of the survey questionnaire and the other of which recommended measures to increase survey response rates.[15] FHFA has carefully considered each of the suggested revisions, but, as explained below, has decided not to implement any of those suggestions at this time.

The first comment letter was from an individual who has served in various capacities with a community association trade group and who is the president of a company that provides online technology in support of the sale, resale, finance, and refinance of homes in community associations.[16] The letter asserts that certain questions in the NSMO questionnaire "fail to adequately and effectively recognize" the role of community associations in U.S. home ownership and that, as a result, data from the NSMO regarding community associations "has nominal heuristic and statistical value at best." It suggests adding several questions to the NSMO questionnaire, and revising several existing questions, to elicit more information relevant to community associations.

Specifically, the letter first suggests revising Question 60 to elicit more specific information on the type of property that is associated with the respondent's mortgage and adding two questions as to whether the respondent's property is in a community association and, if so, the specific type of community association. FHFA believes that, while such questions could be suitable for a survey that focuses on housing structure, they would not be appropriate for the NSMO, which focuses on consumers' experience in seeking and obtaining a residential mortgage loan.[17] The commenter also suggests adding a question to elicit information on the respondent's level of familiarity with various types of community association fees. Again, such a question would be beyond the scope of the NSMO, which does not attempt to capture information on the cost of a mortgage or on fees paid at origination or over the life of the mortgage.

Finally the letter suggests revising the answer choices for Questions 7, 39, and 50 to allow respondents to indicate, respectively: whether they used any of the proceeds from a refinance to pay community association fees; whether and to what extent community association documents or officials may have provided them with information about mortgages or mortgage lenders; and whether and to what extent they sought input about their mortgage loan closing documents from officials of a community association. FHFA notes that each of those questions permits a respondent to choose "other" and to write in a specific answer if none of the other answer choices are applicable. To date, none of the questions have elicited an "other" response in the vein of any of the answer choices that the commenter suggests adding. Accordingly, FHFA does not see a need to revise any of the questions in the manner suggested.

The second comment letter, from a law school professor, states that the NSMO is very important to understanding the health of the mortgage market and agrees that the collection is necessary for the proper performance of FHFA functions. However, it also expresses a concern that, given the

---

[15] Copies of these comment letters are included as Attachment 5.

[16] The letter explains that community associations are "housing management organizations that are an out-growth of traditional subdivision and zoning controls" and include condominiums, cooperatives, and planned communities.

[17] Both the American Housing Survey (sponsored jointly by the Department of Housing and Urban Development and the Census Bureau) and the American Community Survey (sponsored by the Census Bureau) would be more appropriate vehicles for eliciting such information.

length of the survey questionnaire, those recipients who ultimately decide to respond will not be representative of the typical borrower. It suggests two ways of encouraging a response from recipients who might otherwise be reluctant to take the time to complete the survey: (1) providing a greater incentive; and (2) allowing recipients the option of completing a shorter version of the questionnaire.

FHFA agrees that non-response bias (the bias that results when respondents differ systematically from non-respondents) is an important concern and the Agency has spent, and continues to spend, significant time considering ways to increase response rates and to mitigate the effects of non-response bias. In developing the NSMO, the Agency consulted with Dr. Don Dillman and Dr. Mick Couper, two of the foremost experts on conducting consumer surveys. Both experts recommended an up-front payment of five dollars as the most effective way of incentivizing survey recipients to respond and FHFA adopted this recommendation. In addition, based on the results of the Freddie Mac pilot surveys and the first seven waves of the NSMO, Drs. Dillman and Couper also evaluated the expected effect on the response rate of increasing or decreasing the number of questions and the length of the questionnaire. Both experts have opined that shortening the questionnaire would not significantly increase the response rate.

With respect to the mitigation of non-response bias when analyzing survey responses, FHFA has followed best practices of survey sampling analysis. The response to question 3 in part B of this Supporting Statement includes a detailed discussion of the "non-response weighting" method used for the NSMO. As explained there, the availability in the National Mortgage Database of extensive credit and administrative data on both responding and non-responding borrowers gives FHFA the ability to construct non-response weights with more accuracy than is possible for most surveys.

In addition to soliciting public comments, FHFA has sought input on the survey materials through pre-testing. As part of the continuing development of the NSMO, the survey subcontractor, Westat, has conducted cognitive tests of individual borrowers to pretest. The insights gained from such testing assists FHFA in designing the questionnaire that encourages recipients to answer the survey questions and to provide useful data on their experiences with shopping for and taking out mortgage loans. In order to avoid any violation of the PRA, Westat has limited such pre-testing to no more than nine individuals in any year. FHFA is requesting OMB clearance for the pre-testing of the NSMO materials so that cognitive tests may be conducted with 10 or more participants annually while remaining in compliance with the PRA.

9. Provision of payments or gifts to respondents

Survey recipients will receive a cash payment of five dollars as an inducement to complete and return the NSMO questionnaire. Recipients who fail to respond to the first two survey solicitations may receive an additional cash inducement of a similar amount.

Each cognitive pre-testing participant may receive approximately fifty dollars as an incentive payment.

10. Assurance of confidentiality

With respect to the confidentiality of survey responses, the cover letter that accompanies each NSMO questionnaire contains the following statement:

> This survey is voluntary and we ask that you not identify yourself in any way when you return your questionnaire in the enclosed postage-paid return envelope. The code numbers on the survey are there to aid in processing and keep track of returned surveys. No names or other identifying information is ever included in the data.

The questionnaire itself contains a statement, required by the Privacy Act,[18] informing recipients that "[s]ubmission of the survey authorizes FHFA to collect the information provided and to disclose it as set forth" in the current System of Records Notice (SORN) for the National Mortgage Database.[19] The questionnaire also instructs recipients not to include their names or addresses when completing the questionnaire.

Section 1324 of the Safety and Soundness Act authorizes FHFA to modify the mortgage data released to the public as necessary to ensure that it contains no "representation of information that permits the identity of a borrower to which the information relates to be reasonably inferred by either direct or indirect means."[20] For each sampled loan and its associated borrower(s), Experian provides its survey subcontractor, Westat, with the identifying information it needs to administer the survey. However, the data on borrowers and loans that is accessible to FHFA, CFPB, and any other authorized user of the National Mortgage Database, including data obtained through the NSMO, does not include any direct identifying information such as borrowers' names, addresses, or Social Security numbers or the name of any financial institution.

Westat mails a survey questionnaire to the borrower(s) on each sampled mortgage loan at the property address associated with that mortgage. It then uses an encrypted key to track the surveys so that it can compile and maintain the survey opt-out list and identify non-responders to whom it must send follow-up correspondence. All returned questionnaires and any non-delivered mail are sent directly to Westat, not to FHFA, CFPB, or Experian. To maintain the de-identified nature of the data and the confidentiality of the survey responses, Westat purges all responses of any identifying information before providing the collected information to FHFA's National Mortgage Database Project staff for further processing (which is described in Part B of this Supporting Statement).

Similarly, while Westat knows the identity of the cognitive pre-testing participants, that information is not conveyed to FHFA and is not included in the National Mortgage Database in any form.

---

[18] 5 U.S.C. 552a.
[19] *See* 80 FR 52275 (Aug. 28, 2015). A copy of the current SORN for the National Mortgage Database is included as Attachment 6. FHFA is in the process of developing a revised SORN, which it expects to publish in late 2016 or early 2017.
[20] *See* 12 U.S.C. 4544(c)(3), (4).

11. Questions of a sensitive nature

Of the 93 questions on the most recent iteration of the survey questionnaire, approximately 20-25 might be considered to be of a sensitive nature by particular borrowers. Questions that FHFA has identified as potentially sensitive include those requesting information on loan terms, purchase price, household income and other sources of funds, employment status, level of education, age, sex, race, ethnicity, and marital status.

Each of those questions is designed to elicit information that FHFA is required by statute to collect or that is otherwise essential to fulfilling the purposes of the NSMO and the National Mortgage Database Project as a whole. While FHFA understands that some survey recipients will be reluctant to answer questions about these potentially sensitive topics, the Agency believes that others will look upon doing so as an opportunity to express themselves about issues of concern to them.

12. Estimates of the hour burden of the information collection

This information collection comprises two components: (I) conducting the survey; and (II) pre-testing survey questionnaires and related materials through the use of cognitive testing. FHFA estimates that the total annualized hour burden imposed upon members of the public by this information collection will be 12,030 hours: 12,000 hours associated with conducting the survey and 30 hours associated with pre-testing the survey materials. Because the survey recipients and cognitive testing participants are individuals only, there are no hourly costs associated with the burden estimates. The overall burden estimates are based on the following calculations:

I. *Conducting the Survey*

The estimated annualized hour burden associated with conducting the NSMO is 12,000 hours. The NSMO questionnaire will be sent to 6,000 recipients quarterly. Although, based on historical experience, the Agency expects that only 30 to 35 percent of those surveys will be returned, it has assumed that all of the surveys will be returned for purposes of this burden calculation. Based on the reported experience of respondents to prior NSMO questionnaires, FHFA estimates that it will take each respondent 30 minutes to complete the survey, including the gathering of necessary materials to respond to the questions.

Recipients read and complete survey questionnaire and return the completed form to the survey subcontractor:

- Completion time per recipient: 0.5 hours
- Survey mail-outs annually: 4
- Recipients per survey: 6,000
- Total recipients annually: 24,000
- Total hours annually: 12,000 hours

II. Pre-Testing of Survey Materials

The estimated annualized hour burden associated with the pre-testing of the survey materials is 30 hours.

Selected individuals participate in cognitive testing to pre-test the survey questionnaire and related materials:

- Time per participant: 1 hour
- Total participants annually: 30
- Total hours annually: 30 hours

13. Estimated total annualized cost burden to respondents

FHFA has not identified any costs to respondents.

14. Estimated cost to the federal government

The estimated annual burden to the federal government is $480,000 and 400 hours, calculated as follows:

FHFA analyst embeds NSMO data into a query-based electronic database and carries out data cleaning, imputation, and non-response bias weighting:

- Processing time per survey: 100 hours
- Total surveys annually: 4
- Total hours: 400
- Hourly rate: $75 (includes salary, benefits, and overhead)
- Total cost: $30,000

In addition, approximately $450,000 will be paid annually to the contractor hired to conduct the surveys. Of this, approximately $215,000 will be attributable to the cash incentive payments to survey recipients; approximately $60,000 will be for printing costs; and approximately $50,000 will be for postage costs.

$30,000 (hourly cost) + $450,000 (paid to subcontractor) = $480,000.

15. Reasons for change in burden

FHFA has decreased the estimated total annual burden hours to 12,030 from the 14,000 that were shown in the Supporting Statement submitted in connection with the 2013 clearance of this information collection. In 2013, FHFA estimated that the NSMO questionnaire would be sent to 7,000 recipients every quarter. Item #12 in this Supporting Statement reflects the fact that the survey has been, and will continue to be, sent to 6,000 recipients per quarter. This resulted in a decrease of 2,000 hours—from 14,000 to 12,000—in the annual burden associated with the

conduct of the survey. The additional 30 hours in the current burden estimate relate to the pre-testing of the survey questions, for which FHFA did not request clearance in 2013.

16. Plans for tabulation, statistical analysis and publication

On May 27, 2016, FHFA and CFPB released NSMO Technical Report 16-01, the first report on data collected through the NSMO that has been made available to the public. The report is available on FHFA's website[21] and a copy is also included with this Supporting Statement as Attachment 7.

NSMO Technical Report 16-01 presents aggregate data from the 6,062 responses to the first year of the NSMO—specifically, those pertaining to borrowers who obtained a mortgage in 2013 and whose loans were reported to Experian by September 2014. Thus, the report provides an overview of the mortgage market and borrowers' experiences in 2013. These unweighted responses were about one-third of the sample drawn from mortgages originated in 2013. It is intended to be the first in a series of annual reports on new mortgage borrowers based on the NSMO.

The 48 page report is organized into thirteen sections, as follows:

- Section one gives an overview of the NSMO and of the report itself.
- Section two compares the 2013 NSMO data with 2013 HMDA data.
- Section three presents the overall profile of survey respondents in 2013, including their demographic characteristics and what kind of mortgages they obtained.
- Sections four through seven describe how the 2013 borrowers shopped for their mortgages and the application and closing processes.
- Section eight presents borrowers' mortgage outcomes and measures of their satisfaction in the mortgage application and closing process.
- Section nine considers borrowers' opinions on financial responsibility.
- Section ten examines borrowers' opinions on the neighborhood where they obtained their mortgage.
- Sections eleven and twelve discuss 2013 mortgage borrower's expectations for house prices, neighborhood desirability, and their own financial circumstances over the next couple of years.
- Section thirteen describes the anticipated next steps regarding the release of NSMO data.
- The Appendix provides details on the overall National Mortgage Database Project and the sample frame and data preparation procedures for the NSMO.

The report contains 28 tables presenting data in tabulated form. The analytical techniques used in preparing the data for Report 16-01 are discussed in detail in Part B of this Supporting Statement, as well as in the Appendix to the report and in the draft of NSMO Technical Report 2.1 (Attachment 2).

---

[21] *See* https://www.fhfa.gov/PolicyProgramsResearch/Programs/Pages/National-Mortgage-Database.aspx.

Going forward, FHFA and CFPB plan to publish similar NSMO Technical Reports annually, with each presenting NSMO data pertaining to borrowers who obtained a mortgage during a particular calendar year. The agencies are currently preparing the report on data for sampled loans originated in 2014, which is expected to be released in early 2017. The agencies expect to release reports on NSMO data for loans originated in 2015 and subsequent years approximately 24 months after the end of the calendar year to which the data pertains.

17. If seeking approval to not display the expiration date for OMB approval of the information collection, explain the reasons why display would be inappropriate

FHFA will display the expiration date for OMB approval.

18. Explain each exception to the topics of the certification statement identified in "certification for paperwork reduction act submission."

There are no exceptions to the topics of the certification statement identified in the "Certification for Paperwork Reduction Act Submission."

B. COLLECTIONS OF INFORMATION INVOLVING STATISTICAL METHODS

Question 1. Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection methods to be used. Data on the number of entities (e.g., establishments, State and local government units, households, or persons) in the universe covered by the collection and in the corresponding sample are to be provided in tabular form for the universe as a whole and for each of the strata in the proposed sample. Indicate expected response rates for the collection as a whole. If the collection had been conducted previously, include the actual response rate achieved during the last collection.

The NSMO represents a universe of 6 to 9 million recently originated closed-end first-lien mortgage loans on single-family residential property that are reported to Experian annually. A 1-in-20 simple random sample of those mortgages are added to the National Mortgage Database each quarter, resulting in approximately 80,000 to 100,000 loans being added quarterly and 320,000 to 400,000 loans being added annually. For each of the twelve quarterly waves of the NSMO that have been mailed to date, FHFA has selected an approximately 1-in-15 random sample of loans from one or more recent quarterly updates of the National Mortgage Database. This represents an approximately 1-in-300 sample of the universe as a whole.

For each quarterly survey since Wave 8 (mailed in the 4th quarter of 2015), FHFA has randomly selected a sample of approximately 6,000 mortgage loans from those reported to Experian during the preceding quarter and newly added to the National Mortgage Database, with the additional conditions that the loans must have been reported to Experian within a year of origination and that the borrowers must not have been selected for an earlier NSMO survey. The reporting and origination criteria that applied to the selection of sample loans for Waves 1-7 of the NSMO varied somewhat from the criteria that is currently used. Those criteria are shown in Table 1 below.

TABLE 1 Reporting And Origination Criteria Applying to Selection of Sample Loans for NSMO Waves 1-11			
Wave	Mailed	Reported between:	Originated:
1	2014 1Q	4/1/2013 and 12/31/2013	In 2013
2	2014 2Q	1/1/2014 and 3/31/2014	In 2013
3	2014 3Q	4/1/2014 and 6/30/2014	In 2013, but no more than 1 year prior to being reported
3	2014 3Q	1/1/2014 and 6/30/2014	In 2014
4	2014 4Q	7/1/2014 and 9/30/2014	No more than 1 year prior to being reported
5	2015 1Q	10/1/2014 and 12/31/2014	No more than 1 year prior to being reported
6	2015 2Q	1/1/2015 and 3/31/2015	In 2014, but no more than 1 year prior to being reported
7	2015 3Q	1/1/2015 and 3/31/2015	In 2015
7	2015 3Q	4/1/2015 and 6/30/2015	No more than 1 year prior to being reported
7	2015 3Q	1/1/2014 and 6/30/2014	In 2014, but no more than 1 year prior to being reported (these criteria applied to mortgage on properties in "remote rural" counties only)

After the sample is selected, Experian eliminates any potential respondents who have previously opted out of the NSMO[22] or who are deemed not to have legitimate addresses or names.[23] This typically results in somewhat fewer than 6,000 surveys being mailed out in any given quarter. As shown in Table 2 below, the overall response rate for the eleven waves of the NSMO that have been completed to date has been 34.3 percent, with the individual waves achieving response rates of between 30.2 percent and 38.6 percent.

TABLE 2 Size of Universe, Size of Surveyed Sample Size, Number of Responses & Rate of Response for Each NSMO Wave						
NSMO		Loans Meeting Criteria in Table 1		Surveys Mailed & Returned		
Wave	Quarter Mailed	# Reported to Experian	# Added to NMDB	# Mailed	Returned	
					#	%
1	2014 1Q	6,963,160	348,158	15,000	5,793	38.6%
2	2014 2Q	888,420	44,421	3,000	1,076	35.9%
3	2014 3Q	1,685,760	84,288	5,992	2,073	34.6%
4	2014 4Q	1,527,740	76,387	5,795	2,020	34.9%
5	2015 1Q	1,465,380	73,269	5,925	2,089	35.3%
6	2015 2Q	1,057,940	52,897	4,428	1,489	33.6%
7	2015 3Q	2,182,960	109,148	7,359	2,494	33.9%
8	2015 4Q	1,961,820	98,091	5,820	1,875	32.2%
9	2016 1Q	1,753,860	87,693	5,756	1,824	31.7%
10	2016 2Q	1,519,620	75,981	5,788	1,765	30.5%
11	2016 3Q	1,669,620	83,481	5,733	1,733	30.2%
Totals		22,676,280	1,133,914	70,596	24,231	34.3%

Returned survey questionnaires and on-line responses are evaluated to determine the usable responses. Four criteria are used to eliminate unusable responses.

The first criterion is a "no" response to Question 1 ("Within the past 18 months or so, did you take out or co-sign for a mortgage loan including any refinance of an existing mortgage?"). Question 1 is used as a screener question to confirm that the survey respondent took out a mortgage during the reporting period (which Experian records suggest that they did). In Wave 1, a surprisingly high number of respondents—764 out of 5,793—said that they had not taken out a mortgage recently. An analysis of those responses suggested that some respondents who had refinanced their mortgage

---

[22] During the first eight weeks of each wave, Experian maintains a NSMO call center to address any questions by respondents. This call center also allows respondents to opt out of the current survey, as well as any future surveys.
[23] Industry guidance (Metro 2® Industry Standards for Credit Reporting) requires that servicers must supply a billing address for each borrower on all trade lines, including mortgages. Experian generally uses these borrower billing addresses as the survey mailing address.

did not consider the loan to be a new mortgage. Consequently, in Wave 2, the wording of Question 1 was changed to add the phrase "including any refinance of an existing mortgage." With this change, the share of "no" responses to Question 1 decreased from 13 percent to 8 percent.

The second exclusion criterion eliminates breakoffs, defined as those questionnaires for which the respondent did not answer almost all questions from Question 50 through the end, or answered less than 50 percent of the questions overall.

The third criterion excludes questionnaires for which respondents provided information on the wrong loan. The sampling frame is tied to a particular loan associated with the borrower. The questionnaire, however, does not refer explicitly to that loan; instead, it asks respondents who have taken out multiple loans during the reference period to report on the "most recent refinance or new mortgage." In some instances this is not the sample loan. This was a particular problem in Wave 1 which had a longer reference period than subsequent waves.[24] Also, some respondents who have refinanced their mortgage reported on the original home purchase mortgage loan rather than on the refinance. Finally, in some instances it may become apparent that the survey was sent to the wrong person, in that the responses bear no resemblance to the sample loan features as characterized by Experian records. In each of these circumstances, the survey questionnaire is removed from the data set used for analysis.

The fourth criterion excludes questionnaires completed by respondents whose sample loans are removed from the National Mortgage Database after the survey has been executed. Loans may be removed either because they were deemed to have duplicate trade lines (*i.e.*, two or more survey recipients are co-borrowers on the same sampled loan) or because the sample loan has been determined to be a second lien, and not a first lien, mortgage. In some instances, the survey response itself may lead to the removal of the questionnaire under this criterion, if margin notes or comments made by the respondent indicate that the loan was a second lien. This was a particular problem in Wave 4, during which a higher-than-usual number of questionnaires were determined to have related to second mortgage loans and had to be excluded.[25]

The elimination of returned questionnaires meeting any of the above criteria resulted in an overall usable response rate of 30.1 percent for Waves 1-7 of the NSMO (this analysis has not yet been completed for Waves 8-11). Table 3 below summarizes the results of this analysis for Waves 1-7 and shows the number of questionnaires excluded based on each of the four criteria.

---

[24] Wave 1 of the survey, mailed in April 2014, included a sample of 15,000 mortgages. This was a catch up period to cover cases originated in 2013 and newly reported to Experian in the archives for June, September, and December 2013.

[25] At that time, faulty selection criteria had resulted in a significant number of second mortgage loans in the Experian files being pulled into the National Mortgage Database. When this was discovered, these loans were removed and the selection criterion was corrected.

TABLE 3 Usable Responses from NSMO Waves 1-7								
NSMO		Surveys Mailed	Surveys Returned	Answered "No" to Q1	Didn't Finish Survey	Wrong Loan	Duplicate or HELOC	Usable Responses
Wave	Quarter Mailed							
1	2014 1Q	15,000	5,793	748	129	218	57	4,641
2	2014 2Q	3,000	1,076	86	16	40	10	924
3	2014 3Q	5,992	2,073	114	36	65	18	1,840
4	2014 4Q	5,795	2,020	88	46	67	153	1,666
5	2015 1Q	5,925	2,089	86	46	64	26	1,867
6	2015 2Q	4,428	1,489	78	29	53	102	1,227
7	2015 3Q	7,359	2,494	156	75	112	29	2,122
Totals		47,499	17,034	1,356	377	619	395	14,287
Percentages		100.0%	35.9%	2.9%	0.8%	1.3%	0.8%	30.1%

Question 2. Describe the procedures for the collection of information, including:
- **Statistical methodology for stratification and sample selection,**
- **Estimation procedure,**
- **Degree of accuracy needed for the purpose described in the justification,**
- **Unusual problems requiring specialized sampling procedures, and**
- **Any use of periodic (less frequently than annual) data collection cycles to reduce burden.**

Data for the NSMO is collected through a single-blind mail survey format. The survey's sample selection is discussed earlier in Question 1. The NSMO is a simple, random sample of mortgage originations and is not stratified. Alternatives designed for stratifying, clustering, or cut-off samples were not considered. While subpopulations are of interest, the key purpose of the survey is to gather information on the origination characteristics of the sampled borrowers. No estimation procedures are needed in interpreting the survey responses. Similarly, no hypotheses are being tested and no unusual problems exist that require specialized sampling procedures.

Question 3. Describe methods to maximize response rates and to deal with issues of non-response. The accuracy and reliability of information collected must be shown to be adequate for intended uses. For collections based on sampling, a special justification must be provided for any collection that will not yield "reliable" data that can be generalized to the universe studied.

Methods Used to Maximize Response Rates

After achieving a response rate of just 12 percent on its initial pilot survey in 2010, Freddie Mac retained Dr. Don A. Dillman of Washington State University, a leading expert in mail survey

methods, to provide input on ways to maximize response rate. Dr. Dillman provided recommendations in three critical areas: (1) the execution/implementation of the survey; (2) the communications package; and (3) the questionnaire content and format. Freddie Mac adopted Dr. Dillman's recommendations in these areas and, in the second and third pilot surveys, achieved response rates of 60 and 45 percent, respectively. FHFA adopted those recommendations for the NSMO and continues to consult with Dr. Dillman, among others, in developing the survey.

One important recommendation, which FHFA adopted for the NSMO, was to have four planned mail contacts with the survey recipients. The first contact consists of the questionnaire, an upfront five dollar incentive, and a cover letter. The second contact, a reminder letter sent to all recipients, occurs in the second week of implementation. The third contact, in the fifth week, is sent only to non-responders and includes a second reminder letter, another copy of the questionnaire, and a second five dollar incentive. The final contact, a third reminder letter, is sent in the seventh week to non-responders only. A due date for returning the survey questionnaire is included in the last mailing, which closes the communication loop with all survey recipients.[26]

FHFA also adopted two recommendations regarding the communications package. First, all communications have a friendly tone and reflect a personal and sincere request for help. All correspondence is signed by a senior official of both FHFA and CFPB and includes contact information for authenticity. Second, each questionnaire is mailed in a plain white envelope so as not be mistaken for "junk mail."

To further increase response rates, survey recipients are given the option of completing the survey online in either English or Spanish. The first mailing contains an insert, in both English and Spanish, which informs recipients of these options and provides the web addresses to access the appropriate electronic versions of the survey.

Data editing

The survey responses, once delivered to FHFA's National Mortgage Database Project staff, are subjected to thorough editing and review. The initial phase consists of standard data editing—correcting numbers reported in the wrong units, changing answers in responses based on margin notes and comments, assigning responses for questions with open-ended "other" responses, dealing with multiple responses to a question that calls for only one response, and deciding how to handle situations where respondents followed the wrong skip pattern.

Initial examination of the survey responses from Waves 1 through 4 of the NSMO indicated that certain questions may have been misunderstood or misinterpreted more frequently than others. It also appeared that some respondents with sample loans secured by investment property may have provided information on their primary residence property and neighborhood rather than that of their investment property. FHFA addressed these problems by making some changes to the questionnaire for Wave 7 based on the June 2015 archive.

---

[26] Copies of the four letters (in both English and Spanish) comprising the NSMO communications package used for the November 2016 survey mailing are included as Attachment 8.

One advantage that the NSMO has over other surveys is the availability of credit and administrative data that can be used to assist in the editing and imputation processes. The National Mortgage Database draws such data from four primary sources: (1) credit data on the sample loans from Experian; (2) data on the survey respondents collected by Experian from other data sources, including loan servicers and data companies; (3) data from matches to administrative loan and property files; and (4) data from matches to HMDA files. Currently, only matches to Fannie Mae and Freddie Mac administrative files and HMDA data through calendar year 2015 are available and have been used. In the future, additional information from further HMDA and administrative and property file matches will be available for this purpose. The credit and administrative data are used to determine which borrower in the Experian data corresponds to the survey respondent (and spouse/partner of the respondent) and to infer the loan the respondent had in mind when answering the survey. These data are also useful in determining if respondents correctly identified their loan as a home purchase loan or a refinance.

Imputation

After editing and cleaning the survey response data, National Mortgage Database Project staff impute missing responses using statistical models estimated based on credit and administrative data and answers to other questions in the survey. In order to preserve the original responses, the raw responses are retained ("Q" variables) with missing responses coded as such. A parallel set of variables ("X" variables) are constructed where all missing responses are imputed. Each instance in which an X variable differs from its comparable Q variable is recorded by a shadow variable ("J" variables) that indicate the method and reason for which the change was made. Missing responses typically total about 3 to 5 percent for most questions and only in a few instances were more than 10 percent. The X variables are not created when a directly comparable credit or administrative variable is available for all respondents (*e.g.*, loan amount, loan payment, number of co-signers) as comparable credit or administrative variables can be used in lieu of survey responses in analysis.

Key demographic variables (age, gender, education, ethnicity, and income) are imputed first. For these variables, high quality administrative data are generally available and can be used directly to impute a value for the X variable. For example, lender-reported information provides high quality data on age. Administrative data, which are currently available for Fannie Mae and Freddie Mac loans originated in 2013, 2014, and 2015, provide reliable information on race, income, and interest rate. HMDA data, which are currently available for loans originated in 2013, 2014, and 2015, provide reliable information on race, income, and gender.

Comparable credit or administrative information is not available, however, for most survey variables. For these variables, missing values are imputed statistically using an iterative process. More specifically, individual statistical models are developed for each question that used the key demographic variables as well as credit or administrative data such as loan amount and credit score as regressors in linear probability, logistic, or cell-based models (since almost all variables in the survey are categorical). Each imputed value incorporates both a predicted value from the model as well as a random component that reflects the accuracy of the imputation model. Variables are imputed in order, with higher-order variables that dictate a skip-pattern imputed first, before the variables conditioned on the pattern are imputed. Once the first round of imputations is completed,

the process is repeated with expanded models that incorporate some of the newly imputed variables as regressors for other variables. Iterating in this way ensures that correlations among the imputed values will better reflect correlations among observations where responses were available.

Non-Response Weighting

In the NSMO, two types of non-response may occur. The first type is unit non-response, which occurs when data is not obtained for the sample unit—*i.e.*, a respondent chooses not to participate in the survey. The second type is item non-response, which occurs when a respondent fails to answer one or more questions on the survey. To correct for these biases, National Mortgage Database Project staff create non-response adjustment weights and apply them to the survey responses.

Because little is typically known about survey non-responders, the statistical models used to construct non-response weights are quite simplistic. Compared with many other surveys, however, the NSMO has extensive credit and administrative data on both responding and non-responding borrowers that can be used to estimate non-response weights with more accuracy.

Sample non-response weights are estimated with logistic models separately for each sample wave and within a wave for: loans with a single borrower versus those with multiple borrowers; and loans with a HMDA match versus those without a HMDA match. The models estimate the probability of getting a usable response for each wave of the survey. The predictive equations have had pseudo-R-square values ranging from 0.0530 to 0.5103. The largest pseudo-R-square values were for models estimated on data from Wave 7, which had a smaller number of 2014 observations than other waves. This value was notably higher than the pseudo-R-square values for earlier waves. Through Wave 6 the largest pseudo-R-square was 0.2219. Key predictive variables included are: the loan amount; borrower age; median income of borrowers; census tract of the sample loan as captured in the HMDA data; an indicator of investor status; and, if matched to HMDA data, whether it was a home purchase or refinance loan, whether a borrower kept a loan at the same time the sample loan was taken out (an indicator of multiple loans), and a measure of the number of days from loan origination to sending out the survey. The models also control for credit score, for geography using Census Divisions, and for demographic characteristics using Experian's marketing-type variables on family composition, race, ethnicity, gender, and educational attainment.

The model's predicted probabilities of response are grouped into quintiles. The average of the response rates from each of these five groups is used to calculate a response weight as the inverse of these five average rates. Once within-wave sample non-response weights are estimated, they are multiplied by the wave sample weight to provide an overall weight.

Table 4, below, demonstrates the effect of differential sampling weights for Waves 1-7 of the NSMO. Column one shows the distribution among various demographic and loan categories of the raw survey responses. Column two provides the distribution using estimated overall weights. Column three shows the average overall weight for each category.

Table 4 Survey Sample Weights, Waves 1-7 for 2013 and 2014 Originations (Excludes Remote Rural)			
	Unweighted Percentage	Weighted Percentage	Average Weight
Loan Category			
Purchase, First-Time Homeowner	9.0	10.0	1210
Purchase, Repeat Homeowner	33.6	32.1	1035
Purchase, Seasonal Home	1.7	1.3	833
Purchase, Relative Home	0.8	0.6	882
Purchase, Investment Home	2.4	2.3	1033
Refinance, Homeowner Cashout	16.0	15.8	1072
Refinance, Homeowner Regular	32.4	33.8	1130
Refinance, Seasonal Home	0.8	0.7	947
Refinance, Relative Home	0.5	0.4	996
Refinance, Investor Home	2.9	3.0	1127
	100.0	100.0	
Loan Size			
$50,000 or Less	3.7	3.5	1036
$50,001 to $150,000	38.8	39.3	1096
$150,001 to $300,000	38.7	38.4	1077
More than $300,000	18.8	18.8	1083
	100.0	100.0	
Mortgage Term to Maturity			
Less than 15 Years	5.4	4.9	994
15 Years	18.0	17.2	1038
15 to 30 Years	6.0	6.2	1130
30 Years or More	70.7	71.7	1099
	100.0	100.0	
Respondent Credit Score			
Lower than 620	3.6	4.8	1440
620 to 639	3.0	3.7	1345
640 to 659	4.7	5.8	1343
660 to 679	5.3	6.1	1252
680 to 699	6.3	7.0	1202
700 to 719	7.5	8.2	1183
720 to 739	9.3	9.9	1156
740 or Higher	60.4	54.5	980
	100.0	100.0	

Question 4. Describe any tests of procedures or methods to be undertaken. Testing is encouraged as an effective means of refining collections of information to minimize burden and improve utility. Tests must be approved if they call for answers to identical questions from 10 or more respondents. A proposed test or set of tests may be submitted for approval separately or in combination with the main collection of information.

FHFA will use information collected from the cognitive testing participants to assist the Agency in drafting and modifying the survey questions and instructions, as well as the related communications, to read in the way that will be most readily understood by the survey respondents and that will be most likely to elicit usable responses. Such information will also be used to help the Agency decide how best to organize and format the survey questionnaire. A copy of the most recent version of FHFA's NSMO cognitive testing guidance document (or "Talk Track"), which was provided to Westat on October 28, 2016, is included as Attachment 9.

Question 5. Provide the name and telephone number of individuals consulted on statistical aspects of the design and the name of the agency unit, contractor(s), grantee(s), or other person(s) who will actually collect and/or analyze the information for the agency.

The names of and contact information for individual stakeholders from FHFA, CFPB, Freddie Mac, and Experian, including those who were consulted on statistical aspects of the design and who will analyze the data, appear in the list included as Attachment 10. FHFA also consulted with the following:

Dr. Mick P. Couper
Survey Research Center and the Institute for Social Research
University of Michigan
426 Thompson Street
Ann Arbor, MI 48104
(734) 647-3577

Dr. Don A. Dillman
Department of Sociology and the Social & Economic Sciences Research Center
Washington State University
Pullman, WA 99164-4014
(509) 335-1511

The subcontractor hired by Experian to carry out the survey and the cognitive testing is:

Westat
1600 Research Blvd,
Rockville, MD 20850