BUREAU OF CONSUMER FINANCIAL PROTECTION
PAPERWORK REDUCTION ACT SUBMISSION
INFORMATION COLLECTION REQUEST

SUPPORTING STATEMENT PART B COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

OWNING A HOME EVALUATION STUDY
(OMB CONTROL NUMBER: 3170-XXXX)

Describe (including a numerical estimate) the potential respondent universe and any sampling or other respondent selection method 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 has been conducted previously, include the actual response rate achieved during the last collection.

The target audience for the Owning a Home suite of tools is prospective homebuyers. To replicate this target audience as much as possible, the respondent universe is composed of qualified users of a national home buying website. Qualified users are defined as those who have registered accounts and who have both opted-in to receive marketing emails and have saved criteria for a home search on their website profile. A random sample of qualified users will be invited to participate in the study via email and directed to visit the study's homepage. The partner website will send these recruiting emails using their email management system. We will use phased recruitment waves to target an initial pool of approximately 230,000 potential respondents who receive the email and click through to the study homepage. We may adjust this target up or down based on early results from the initial recruitment waves.

We acknowledge that qualified users may be different from the general population of prospective homebuyers, especially insofar as they are likely to be especially comfortable with online tools and resources. However, in order to recruit consumers who are actively shopping for a home, we need access to a large number of people who can easily be identified as prospective homebuyers. We do not intend to extrapolate our findings to larger populations, including a general population of internet users or homebuyers. Nevertheless, based on results from our pilot study, we expect that the final sample will be diverse in terms of age (25-29 to over 65) and credit scores (less than 580 to over 760). Respondents from our pilot study varied in terms of marital status (62.6% married), education (5% high school graduates, 28% with an associate's degree or some college, 39% college graduates, and 29% with postgraduate studies), gender (% male), yearly income (8% less than $35,000, 6% between $35,000 and $49,999, 23% between $50,000 and $74,999, 20% between $75,000 and $99,999, 24% between $100,000 and $174,999, and 19% with $175,000 or more), and other characteristics.

Potential respondents will be asked to complete a short screening questionnaire to determine eligibility for the study. To be eligible, a respondent must intend to purchase a home in the next 3 months, intend to pay for that purchase using a mortgage, be involved in financial decisions in his/her household, and not be professionally involved in the real estate industry.

Based on pilot experience, we anticipate that approximately 74% of the potential respondents who arrive at the study home page will complete the screener, of which about 25% will be found eligible. Of those eligible respondents, we anticipate about 59% will agree to participate in the study, for an overall study pool of approximately 25,000 and a participation rate of 11% of the recruited pool.

Table 1. Survey sample and estimates of response rate

Qualified members of homebuying site	10,000,000 +
Recruited pool: invited respondents who arrive at the study homepage	230,000
Respondents who complete the screener (74%)	170,200
Respondents who are eligible (25%)	42,550
Eligible respondents who opt-in (59%)	25,105
Participants who complete the study (16% of opt-ins; 1.8% of recruited pool)	4,122

As this is a longitudinal study occurring over three months, there is likely to be considerable attrition over time. Based on pilot experience, we estimate that approximately 4,122 participants will complete the entire study, amounting to approximately 16% of the participants who opted in and approximately 1.8% of the overall recruited pool. Monetary incentives will be offered to encourage participation in the study and to incentivize survey completion. Further efforts to increase response rates are detailed below.

1. Describe the procedures for the collection of information:

Data will consist of a combination of participant surveys, administrative data from the Owning a Home suite of tools, and data from participants' mortgage documents. Before analysis, data will be de-identified to exclude any direct identifying information protect the privacy of sampled consumers.

Participant surveys: Surveys will be used to obtain additional information on respondents' shopping behavior, perceptions of the mortgage market, expectations of personal mortgage terms, mortgage knowledge, and feelings of empowerment. Surveys will be administered online every two weeks for three months or until a home is

purchased, whichever occurs first. An initial baseline survey will additionally contain information on respondents' background characteristics, including an overview of their existing assets and liabilities. Closing surveys will ask for information on the respondent's experiences at closing, the terms of their home purchase and the mortgage that they obtained, including a request to provide the final mortgage documents. If participants have not purchased a home before the end of the three month study window, they will be asked to complete a final closing survey, administered upon closing and up to 8 weeks later.

Owning a Home web suite of tools: Website metrics will include measures of participants' activity on the site. For instance, participants may access various tools (e.g. background information on the process of searching for a mortgage, a mortgage rate benchmarking tool, or a guide to loan options), and may spend varying amounts of times on each. To ensure participants' privacy and eliminate re-identification risk, the CFPB will provide web analytics data to the study contractor using a set of unique participant ID codes provided by the contractor through the referral links. The contractor will match the web analytics data with the rest of the study data using this ID code. The contractor will then assign a new, different, and randomized unique ID code to each record after matching the data. The CFPB will receive only this second identifier in the final de-identified analysis dataset.

Mortgage document data: Mortgage data will be provided either through direct access to mortgage documents or by respondents answering survey questions about their mortgage terms, whichever method participants prefer. Specifically, respondents who provide their documents directly have the option to upload an electronic document (including taking a picture of the forms) or faxing the information. Those who choose to answer survey questions will be asked basic mortgage information, such as the amount of the loan and the interest rate. Survey questions will be administered during standard survey procedures.

a. Include statistical methodology for stratification and sample selection,

As described above, the sample will be drawn from a home-buying website's existing registered users. Potential respondents will be asked to complete a short screening questionnaire first in order to determine if they are eligible. Data obtained in the screening questionnaire does not contain any directly identifying information and will be analyzed only to assess eligibility. Upon completing the screening questionnaire, respondents are presented with informed consent language and asked to opt-in to the study.

The research team does not anticipate the need to rely on stratification, although respondents may be analyzed according to individual characteristics (e.g. previous homeownership

experience). Findings from this study will not be generalized to larger populations, including internet users, all homebuyers, or the U.S. population.

b. Estimation procedure,

The primary analytic technique for estimating the impact of a randomized controlled trial (RCT) such as this study is a comparison of outcomes for participants in different study groups. In this study, the research team will compare outcomes for participants who were given access to the Owning a Home tools (the first treatment group) to outcomes for participants who were encouraged to shop for their mortgage (the second treatment group) and to those who did not receive any additional resources (the control group). The research team intends to use "intent to treat" analysis, thereby estimating effects on those who had the opportunity to receive treatment, including those who did and did not use the tools. This analysis will identify the causal impact of the treatments on mortgage outcomes for those in the study population.

Estimates of the differences between the two groups will be calculated using data from surveys, usage of the Owning a Home website, and mortgage outcomes. For each difference, statistical confidence intervals will be calculated using standard statistical assumptions.

For some outcomes, the research team is additionally interested in how outcomes changed during the home shopping process. For example, it may be expected that respondents who have completed their home purchase and have closed on a mortgage will have higher knowledge of the mortgage process than those who are in the initial stages of their home search. As such, changes in mortgage knowledge between baseline and the final survey will be compared for those in treatment and control groups.

c. Degree of accuracy needed for the purpose described in the justification,

There is limited existing data on the methods that prospective homebuyers use to search for home and mortgage options. Additionally, there is limited evidence regarding how much money homebuyers save from obtaining an additional mortgage estimate. As such, it is difficult to estimate the expected effects of the described intervention, and in turn, the degree of accuracy required.

Based on power analyses using existing information on estimated savings from mortgage search, the research team has estimated that approximately 3,300 participants are needed to complete the study in order to detect a difference between those exposed to the Owning a Home tools, those who receive the shopping treatment, and those who receive neither. Given uncertainty around the level of attrition in the later stages of the study, the recruitment plan is sufficient to produce 3,300 complete participants with some additional buffer.

d. Unusual problems requiring specialized sampling procedures, and

The research team does not anticipate using any specialized sampling procedures.

e. Any use of periodic (less frequent than annual) data collection cycles to reduce burden.

The proposed research is a one-time collection (i.e., it will not be repeated annually), however the collection is accomplished using several discrete surveys. Respondents will be asked to report information every two weeks, for a period of approximately three months. This frequency of survey was chosen so as to ensure that we can capture the changes in consumers' attitudes and behaviors over a relatively short mortgage shopping/acquisition window. To minimize burden, the individual surveys will be kept short.

2. 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.

Non-response may be a challenge for this study, particularly because respondents are being asked to participate multiple times throughout their shopping process. While we will strive to minimize attrition over the course of the study, we will also include data from participants who partially participated where appropriate. The survey modules have been carefully designed to collect the data needed for analysis while minimizing burden to participants. Additionally, a combination of strategies is being employed to address non-response, as outlined below.

Strategies to maximize participation

a. Participants will be offered a monetary incentive, which has been shown to increase response rates.[1]
b. Reminder emails will be sent to those who have not taken a survey or who dropped out in the middle. These emails will request that participants complete the survey. Such reminders have been shown to increase response rates.[2]

Strategies to minimize incomplete responses

[1] See Alexander, Devine, Couper, McClure, Stopponi, Fortman, Tolsma, Strecher and Johnson (2008) "Effect of Incentives and Mailing Features on Online Health Program Enrollment," *American Journal of Preventative Medicine*, 34(5): 382-388.
[2] Crawford, Scott D., Mick P. Couper, and Mark J. Lamias. (2001). "Web Surveys: Perceptions of Burden," *Social Science Computer Review, 19*(2): 146-62; Cook, Colleen, Fred Heath, and Russel L. Thompson. (2000). "A Meta-Analysis of Response Rates in Web- or Internet-Based Surveys," *Educational and Psychological Measurement, 60*(6): 821-36.

a. To reduce respondent burden, the survey has been constructed so that participants will answer only a subset of questions for each survey. Specifically, the survey modules that are administered will change depending on a small set of initial questions that allow us to quickly assess the respondent's current situation. By reducing the number of questions that are asked (and the associated time required to respond) more respondents are expected to complete the surveys.
b. Additionally, participants will be allowed to "suspend" their survey session so that they can "resume" it at a later date. Such a design accommodates technical problems that may prohibit participants from completing the survey, such as an Internet connection issue, as well as scheduling constraints.[3]

3. Describe any tests of procedures or methods to be undertaken.

In preparing for this study, the CFPB has tested the survey instrument using cognitive interviews and tested the data collection procedures in a pilot study.

The CFPB conducted two sets of cognitive interviews in order to refine the survey instrument.[4] The first set of interviews was conducted on September 25 and 26, 2014 while the second set was conducted from March 6, 2015 through March 10, 2015. During both sets of interviews, participants were asked open-ended questions regarding their understanding of study materials, revealing problems with question comprehensibility and response options. Study materials were subsequently refined to address these issues.

The pilot study was conducted from October 27, 2014 through November 9, 2014. This study was conducted using procedures that were designed to parallel those from the full study and was intended to assess only methodological issues. Specifically, participants were recruited through an email solicitation, answered screening questions, and completed the baseline survey and up to two periodic surveys. No participants closed on a home during this pilot, and therefore we did not receive any data surrounding mortgage outcomes or changes in mortgage knowledge/empowerment from baseline to closing.

We assessed the pilot study data for item-level non-response and participant attrition, concluding that we should try to reduce respondent burden as much as possible. As such, the CFPB revised the survey instrument and implementation method. An example of each change is as follows:

- **Survey instrument**: Questions with low variance were eliminated in order to reduce the length of the survey instrument. For example, 83.0% of respondents in the pilot study could correctly identify the concept of "home equity"; therefore, we no longer ask this

[3] This practice is recommended by survey methodologists; see pages 336-339 of Couper, Mick P. (2008). "Designing Effective Web Surveys," New York, NY: Cambridge University Press.
[4] For additional information on cognitive interviewing, see Willis, Gordon B. (2005). "Cognitive Interviewing: A Tool for Improving Questionnaire Design," Thousand Oaks, CA: Sage.

question. Eliminating questions reduces the length of the survey instrument and corresponding respondent burden.

- **Implementation method**: The survey invitation now contains unique links so that participants do not have to enter a survey password. This change reduces respondent burden and is expected to reduce attrition.

4. **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.**

Lead researcher
Dustin Beckett
Economist
Division of Research, Markets and Regulations
Consumer Financial Protection Bureau
202-435-9399

Data collection project lead
Ricardo Carvalho
Senior Researcher
Fors Marsh Group
703-598-6046

Survey methodology consultant
Mick Couper
Research Professor
Survey Research Center, University of Michigan
734-647-3577