Justification of Non-material Change

December 8, 2014

(0938-1221)

We request approval of the beta test component of the Marketplace Survey with the following revisions:

1. Update to Questionnaire
2. Update to Survey Administration Mode
3. Update to Survey Administration Language
4. Update to Sampling Design
5. Update to Burden Estimates (Hours and Wages)

These updates are based on information from the Marketplace Survey psychometric test (June to August 2014) and psychometric analyses (September to December 2014). The Marketplace Survey beta test procedures and materials must be completed and approved to enable beta test activities scheduled to begin February 16, 2015.

1. Questionnaire Revisions

Overview of Marketplace Survey Revisions for Beta Test

Overall, we dropped 15 questions to reduce the length of the survey and added 2 new questions on re-enrollees and 1 new question on multiple chronic conditions. There is a net decrease of 12 questions which brings the total survey questions from **95** to **83**. The most substantial changes to the Marketplace Survey for the Beta Test have to do with revisions or restructuring to address the experiences of re-enrollees. We defined re-enrollees as those who had health insurance through the Marketplace last year rather than just interacted with the Marketplace. We use this definition because the process of updating family and income information and re-selecting a health plan only applies to people who already enrolled in a plan. A summary of all the changes to the Marketplace Survey for the Beta Test are listed below followed by more detailed descriptions of the most substantive changes.

- Changed the reference period from October 1, 2013 to November 15, 2014, to align with the start of open enrollment in 2014
- Dropped 15 questions to reduce length (more detail below)
- Added a new question that measures multiple chronic conditions (more detail below)
- Added a new first question in the survey that identifies re-enrollees(more detail below)
- Added a question in the '*Choosing a Health Plan*' section about whether re-enrollees chose the same health plan they had in 2014 through the Marketplace (more detail below)
- Added a new response option about "not finding the same health plan you had in 2014" in the information seeking questions that ask about the reasons why someone did not get the information they needed from the website, phone, or in-person to account for re-enrollee experiences (more detail below).
- Added 'or update' into questions that asked about giving information about the people in your family and giving household income information to the Marketplace to account for re-enrollee experiences (more detail below).

- Reworded 'the people in your family, including yourself' to say 'yourself or the people in your family' to avoid the problem of individuals skipping out of questions because they were not giving information about other family members to the Marketplace.
- Reworded 'What kind of information was not easy to understand' to 'What kind of information was hard to understand' to make the question cognitively easier to comprehend, especially during a telephone interview.
- Reworded 'customer service Help Line' to 'customer service Call Center' to align more closely to the terminology used by Marketplaces.

Detailed Descriptions of Marketplace Survey Revisions for Beta Test

Drop Questions. We recommend dropping 15 questions in an effort to shorten the survey:

1. Q2: Were any of the following a reason why you did **not** give information about the people in your family, including yourself, who wanted health insurance? *Mark one or more.*
2. Q7: Were any of the following a reason why you did **not** give your household income information? *Mark one or more.*
3. Q10: How did you give your household income information?
4. Q14: Since October 1st, were you told by the {INSERT MARKETPLACE NAME} how to appeal the decision?
5. Q15: Was it easy to understand how to appeal the decision?
6. Q32: Since October 1st, how often did the {INSERT MARKETPLACE NAME} customer service Help Line use words or phrases you did not understand when you called?
7. Q37: Since October 1st, did you want in-person help but were unable to get it because the building was not accessible for persons with disabilities?
8. Q43: Since October 1st, how often did the persons you met with about getting health insurance from the {INSERT MARKETPLACE NAME} use words or phrases you did not understand?
9. Q70: Since October 1st, did you get health care 3 or more times for the same condition or problem?
10. Q71: Is this a condition or problem that has lasted for at least 3 months? Do not include pregnancy or menopause.
11. Q72: Do you now need or take medicine prescribed by a doctor? Do not include birth control.
12. Q73: Is this medicine to treat a condition that has lasted for at least 3 months? Do not include pregnancy or menopause.
13. Q87: Are you eligible to get health services from an Indian Health Service, tribal, or urban Indian health program?
14. Q88: Did you ever get health services from an Indian Health Service, tribal, or urban Indian health program?
15. Q93: Do you feel comfortable using the internet through a computer, tablet, or smart phone?

Questions 2 and 7 were dropped because they were designed to capture tourists or people who were exploring the Marketplace but never intended to purchase health insurance, which we believe will not be as common after the first year of open enrollment. These questions created

very complicated skip patterns in the mail survey (more than 30% did not follow the skip pattern correctly). Also, only 11–12% of people actually answered these questions.

Question 10 was dropped because the information was redundant with Question 5 since most people gave information about the people in their family and their income information using the same mode (web, mail, phone, or in-person).

Questions 14 and 15 were dropped because of low item response. Q15 was an assessment item that had to be dropped in our psychometrics due to low covariance coverage with other items. Q14 is the screener for Q15 so it was removed as well. We recommend keeping Q13 which asks if they were told they could appeal a decision about how much they had to pay for their health insurance because we think it is important to track the percentage of enrollees who received this information over time.

Questions 32 and 43 performed poorly in the psychometric analyses. They had low correlations with their scales and lacked discriminant validity (they cross-loaded with other scales). The internal reliability (Cronbach Alpha) and inter-unit reliability for their scales both improved if these questions were dropped.

Q37 was dropped because it had both complicated skip patterns and a low screen-in rate of 2%.

Q70–Q73 were dropped in favor of one question intended to identify multiple chronic conditions. Q70–Q73 measure chronic condition status without identifying how many chronic conditions the respondent has. In an effort to reduce the length of the survey and focus on multiple chronic conditions, which is a more important issue for policy and oversight purposes, we dropped the four CAHPS questions that measure chronic condition status and wrote a new question that measures the presence of multiple chronic conditions. Q87–88 measure eligibility and utilization of Indian Health Services. Less than 1% of psychometric test respondents screened into these questions. We do not believe these questions provide useful information regarding Native American experiences with the Marketplace. With a large enough sample size we could still measure Native American experiences with the Marketplace by using the self-identification as Native American from the race question.

Q93 about comfort using the internet was dropped because it is correlated with age and the relationship between being comfortable using the internet .and Tthe website global rating disappears when age is added to regression model.

New Question on Multiple Chronic Conditions. We believe measuring chronic conditions is important and could affect Marketplace experiences. In an effort to reduce the length of the survey and focus on multiple chronic conditions for policy and oversight purposes, we dropped the four CAHPS questions that measure chronic condition status and wrote a new question that measures multiple chronic conditions. The new question is, "In the last 12 months, did you get care for 2 or more health problems or conditions that each lasted for at least a year?"

New Question to Identify Re-enrollees. We wanted to add a question in the survey that could distinguish re-enrollees from new enrollees in order to do analyses where we compare their experiences. The new question is: "Did you have health insurance through the {INSERT MARKETPLACE NAME} at any time in 2014?" Yes/No. We defined re-enrollees as those who had health insurance through the Marketplace last year rather than just interacted with the Marketplace. This is because the process of updating family and income information and re-selecting a health plan only applies to people who already enrolled in a plan.

New Questions about Difficulty Finding Same Plan as Last Year. Our Technical Expert Panel suggested adding a new question about how easy it was for a re-enrollee to find their same health plan from last year. The assumption is that it may not be that easy to do. Re-enrollees may not remember the marketing name of their health plan or they may have trouble entering in the long ID number associated with their plan. The benefit or cost structure may have changed so re-enrollees may not think it is the same health plan. The health plan name may have changed or may not exist anymore.

We already have a question that measures difficulties with choosing a plan, "was it easy to choose a health plan" that can apply to re-enrollees and new enrollees. In addition, we decided to add a new response option within the questions that ask about "reasons why someone did not get the information they needed from the website, phone, in-person" to address this issue more specifically. The new response option would be: "You/They could not find the same health plan you had in 2014." To measure whether someone does not know if they were enrolled in the same plan last year we ask a follow-up question to those who say they chose a plan during open enrollment for 2015 coverage, "Were you enrolled in that health plan in 2014?"

Modify Existing Questions for Re-enrollees. We modified existing questions to address the experiences of re-enrollees who went back into the Marketplace to update their family and income information. For example we changed "did you *give information* about the people in your family, including yourself, who wanted health insurance through the {INSERT MARKETPLACE NAME}?" to "did you *give or update information* about yourself or the people in your family who wanted health insurance through the {INSERT MARKETPLACE NAME}?" We know some people will only verify their information and not make any changes, but it seems that 'update' is the word being used with consumer facing materials from the Marketplace.

2. Survey administration: Mode

In Part B 'Section 2: Information Collection Procedures' we assumed that the Marketplace beta test would be conducted using a mixed mode design of mail with phone follow-up unless the mode experiment in the psychometric test suggested otherwise. There were 5 experimental mode conditions tested in the English sample only including, mail with phone follow-up, mail with FedEx follow-up, mail with all first class, phone only, and web only. Response rate results from the Marketplace Survey psychometric test mode experiment show that the mail with phone follow-up mode produced the highest response rates (32%) compared to mail with FedEx follow-up (27%), mail with all first class (20%), phone only (22%), and web only (9%) (see exhibit 1).

Mail with phone follow-up and mail with FedEx follow-up were the two modes with the highest response rates, but mail with phone follow-up was less expensive on a cost per complete basis. The overall cost per complete for English cases in the mail with FedEx nonresponse follow-up group was approximately $75 compared to approximately $60 for English cases in the mail with phone nonresponse follow-up group.

There is also evidence from the psychometric test that there are differences in the characteristics of people who respond by phone, mail, or web. The phone respondent distribution is generally more comparable to the frame distribution. The phone captures a higher percentage of potential enrollees and participants who are younger, Black, have lower incomes, and who have a disability compared to mail and web. The mail captures a higher percentage of participants who are older and APTC or CSR eligible compared to phone and web. The web was not as successful

as mail and phone in bringing in a diverse group of respondents. Including phone follow-up will allow us to capture a more diverse population and make our sample more representative of the frame.

For all of these reasons, we plan to administer the Marketplace Survey using a mail with phone follow-up mode in the beta test for all languages. For the English sample only, we will continue to offer the online mode in the pre-notification letter as a way to get early responses.

Exhibit 1. Marketplace Survey psychometric test response rate by mode

Mode	Total completes	Final Response Rate
Mail TOTAL	2,098	28%
Mail w phone follow-up	473	32%
Mail Fed-Ex follow-up	403	27%
Mail all first class-English	301	20%
Mail all first class-Spanish	381	25%
Mail all first class-Chinese	540	36%
Web only TOTAL	133	9%
Phone only TOTAL	323	22%
TOTAL ALL MODES	2,554	24%

3. Survey administration: Language

In Part B Section '1.1.2.2 Marketplace Beta Test Phase Sample Size Estimates' we stated, "CMS will not distribute the survey in Spanish and Chinese, but will administer Spanish and Chinese versions to respondents that request surveys in these languages." However, based on the psychometric test results we decided that we will distribute the beta test survey in Spanish and Chinese to those who identify Spanish or Chinese as their preferred language. We successfully implemented this strategy in the psychometric test and found higher response rates for Chinese (36%) and Spanish (25%) compared to English (20%) when using the same mail with first class mode. Our survey vendor, Ipsos, has found similar results for other health surveys.

Accommodation theory[1] suggests that mailing the surveys in Spanish or Chinese demonstrates cultural sensitivity that leads to positive affect and an increased willingness to respond. It also makes it easier to respond by avoiding the extra effort of calling to request a survey in Spanish or Chinese and the delay in waiting for the survey to be mailed. Prior experiments have also shown that mailing survey materials in Spanish increases response rates over sending them in English only[2].

---

[1] Koslow, Scott, Shamdasani, and Touchstone. 1994. "Exploring Language Effects in Ethnic Advertising: A Sociolinguistic Perspective." Journal of Consumer Research 20:575–85.

[2] Brick, Montaquila, Han, and Williams. Improving Response Rates for Spanish Speakers in Two-Phase Mail Surveys. Public Opinion Quarterly 2012 76: 721-732.

4. Changes to Sampling Design

Overview of Marketplace Sampling Revisions for Beta Test

In Part B of the supporting statement, CMS proposed a sampling design based on assumptions made prior to the psychometric test data collection and analysis. Since ~~this~~ that time, CMS has conducted a detailed analysis of the psychometric test data. In general, these analyses provide information that supports the proposed design aimed at obtaining 1,200 completed surveys from each state participating in the beta test. However, some findings, along with a change in the Marketplace landscape, require two general modifications related to the sampling design.

The proposed modifications include:

1. A change in the number of states participating in the beta test.

2. A change in our approach to sampling consumers whose language preference is Chinese.

Changes in Number of States

In Parts A and B of the supporting statement referenced above, sample size estimates and burden were based on the assumption that all 50 states and the District of Columbia would participate in the beta test. As of December 2014, the number of states expected to be included in the beta test is 44, and this number could decrease (it will not increase).

The decrease in the number of states is due to several factors:

1. CMS has not made participation in the beta test mandatory for the State-based Marketplaces (SBMs); however, seven current SBM states (CA, CT, HI, KY, MN, RI, WA) have indicated that they will voluntarily participate in the beta test.

2. Two states—Oregon and Nevada—are shifting from the State-based Marketplace (SBM) pool of states to the Supported State-based Marketplace (SSBM) pool. Both the SSBM and FFM pools use Healthcare.gov and thus these two states will be part of the sampling frame used for the beta test.

3. Conversely, Idaho is moving from being a SSBM to the SBM pool and is not among the SBMs participating in the beta test.

4. These shifts result in a total of 44 states participating in the beta test: 27 FFM states, seven State Partnership Marketplace (SPM) states, three SSBM states, and seven SBM states.

Given the decrease in the number of states, the total sample size will decrease, as will our estimate of the total number of completed surveys and of the burden. In the sections below we describe some revised assumption about response rates at survey-level and item-level—both overall and by language—and the impact of these assumptions on sample size.

Sampling based on Language Preference

As described in Part A of the supporting statement, one goal of the psychometric test was to evaluate the equivalence of measurement properties across language (see Exhibit A1 in Part A of the supporting statement; see also Section 1.1.2.1 of Part B). Due to a lower than expected survey response rate (RR) among the Spanish language segment of the sample (we had assumed 30% but the actual RR was 25%), combined with low item-level RR for some sections of the

survey, CMS was not able to obtain enough completed surveys in Spanish (n=318) and Chinese (n=540) during the psychometric test to fully evaluate the equivalence of measurement properties across language for the full set of composites. Even the higher survey response rate among Chinese was not enough to counter-balance the low item-level response rates to some survey questions. Our sample size estimates for the psychometric test were based on an assumption that item-level response rates would average around 67% (see section 1.1.2.1 of Part B); in practice, item-level response rates across all languages were as low as 28% on average for some items.[3]

This rate varied substantially by language. For example, for the 'Seeking Information In-Person' section of the survey, the item-level response rates were approximately 20%, 50%, and 32% for English, Spanish, and Chinese respectively, which corresponds to a weighted average of 28% across all three languages (English respondents make up the majority of respondents, so their item-level response rate contributes disproportionately to the average). Thus, even though we had obtained 540 surveys completed in Chinese, only around 170 of those consumers responded to each of the questions about seeking information in-person (q37 to q45 on the psychometric test version of the survey).

Exhibit 2 displays the item-level response rates for the five composites that emerged from the psychometric test analysis, both overall and by language.

Exhibit 2. Response rates for final composites by language

Composite	Overall	English	Spanish	Chinese
Application Process	95%	95%	93%	95%
Seeking Info on Website	63%	70%	44%	55%
Seeking Info over the Phone	53%	53%	59%	50%
Seeking Info In-Person	28%	20%	52%	33%
Health Plan Enrollment Process	85%	86%	82%	85%

To rectify this problem, CMS may need to oversample those consumers with a Chinese language preference. We cannot say definitively at this time whether oversampling is needed because of uncertainty about how many Chinese language consumers will be available in the sampling frame for the beta test states.

Number of Chinese Surveys Needed for Psychometrics. The beta test version of the survey has 26 assessment items. We would thus need 260 completed surveys (10 per item) to conduct the analysis of the equivalence of measurement properties across the three languages. However, since composite-level and item-level response rates are all lower than 100%, we need to inflate

---

[3] The item-level response rate indicates the proportion of respondents who provided a usable response to a given item or composite. Composite-level response rates can be higher than item item-level response rates because a composite is scored for any respondent who provided a response to at least one item in the composite. In section 1.1.2.1 of Part B, we indicated that we needed a minimum of 10 complete responses for each assessment item that was to be used in the psychometric analysis. We inflated this number to 15 based on assuming an average item-level response rate of 67% to account for missing data (10/0.67 = 15). Responses can be missing for two reasons: 1) the respondent skipped the item because it was not applicable to them based on their response to a screener question (legitimate skip), and 2) the respondent did not provide a response to the question, even though it applied to them (non-response). The observed rate of legitimate skips in the psychometric test was higher than we had assumed it would be.

the number of completed surveys enough to ensure 260 completes for the items with the lowest response rate. As described above, the 'Seeking Info In-Person' items yielded an overall response rate of 28%, though the Chinese respondents had a slightly higher rate of 33%. Applying this item-level response rate to the minimum number needed (260), we would need approximately 788 completed surveys in Chinese to conduct this analysis (260/0.33 = 787.9).

Number of Expected Completes from Existing Sampling Design. The next issue is determining if a sample of 1,200 per state across all 45 states would naturally provide enough completed surveys in Chinese and Spanish to meet this goal. Data from the psychometric test indicate that, for the 36 FFM states included in the psychometric test, consumers with a Spanish language preference comprise around 5% of the total survey-eligible population in those states, while consumers with a Chinese language preference comprise around 0.20% of the total survey-eligible population. These estimates, however, are based on a pool of states that exclude those states currently expected to be in our beta test—such as CA, HI, WA, and MD—where the proportion of consumers with a preference for Chinese may be much higher than we have observed in the pool of FFM states. The same can be said for Spanish with respect to CA, and perhaps other SBM states as well.

CMS had estimated before the psychometric test that consumers with a Spanish language preference would comprise around 10% of the frame and that consumers with a Chinese language preference would comprise around 2%, assuming that the sample would include the entire nation (50 states and D.C.). We believe that the estimates obtained from the psychometric test data are more realistic so we will use those instead of our original estimates.

Exhibit 3 displays the total number of completed surveys we can expect to get by language taking into consideration population share in combination with survey response rates observed for these sub-populations in the psychometric test. As shown, we will obtain a sufficient number of completed surveys in Spanish. For Chinese, the number of completes will be insufficient without oversampling.

Exhibit 3. Estimated number of completed surveys by language – Marketplace Survey beta test

Inputs	Overall	English	Spanish	Chinese
Share of population	100%	95%	5%	0.20%
Estimated RR	32%	32%	25%	36%
Number sampled	167,273	156,420	10,560	293
Number completes	52,800	50,054	2,640	106

We will not know with certainty what the total size of the Chinese preference target population will be until we have constructed the sampling frame for the beta test in late February or early March 2015. However, our estimates from the psychometric test suggest it is likely that the target population share of consumers with a Chinese language preference in the frame will be less than 1.5% (given the assumptions laid out above, a population share of 1.5% would yield around 788 completed surveys in Chinese); therefore, CMS will need to oversample this population.

Oversampling will be conducted using a strategy similar to what was used for the psychometric test:

1. CMS will create a separate sampling frame that includes only consumers who indicate a Chinese language preference; this population will thus comprise its own separate stratum.

2. CMS will sort the Chinese frame by state and a random number and draw a systematic random sample of 2,189 consumers from the frame. With this implicit stratification by state, the size of the sample drawn from each state will be proportional to the population share of consumers with a Chinese language preference in that state.

3. Assuming a 36% response rate, a sample of 2,189 will yield 788 completed surveys in Chinese.

Summary of Changes to Sampling Design

In sum, we propose the following changes to the design:

5. The number of states will decrease compared to the original design from 51 to 45.

6. The total target sample size *per state* will not change; it remains 1,200 per state, but because of the decrease in the number of states, the total completed surveys will decrease from 61,200 to 52,800.

7. Sample sizes will change based on observed survey response rates and the validity of assumptions about the distribution of consumers across language preference.

    1. Observed psychometric test response rates for the proposed mail with phone follow-up mode of survey data collection are consistent with our original assumption (32% compared to 30%).

    2. Observed item-level response rates varied considerably across sections of the survey: from a low of 28% to a high of 95%. The design in Part B had assumed 67%. The variance in response rates for the items is reflected in the composite-level response rates. These rates also varied by language, especially for the 'Seeking Info In-Person' composite.

    3. CMS needs a sufficient number of completed surveys in non-English languages to complete the psychometric analysis of the survey measures: 788 in each language.

    4. Sample size estimates must take into consideration both observed RRs and observed item-level response rates, and do so separately by language.

    5. Compared to what was proposed in Part B, the total sample size per state (1,200) will not change; but the distribution of the sample by language will.

    6. There should be more than enough consumers with a Spanish language preference in the frame to produce the requisite number of completes for the psychometric analyses; therefore oversampling is not needed.

    7. Consumers with a Chinese language preference will most likely need to be oversampled.

5. Updates to Burden Estimates (Hours & Wages)

The estimated burden for the Marketplace Survey beta test is shown in Exhibit 4.

Units. For the Beta Test, respondents will be sampled from each individual state, including states with a State-based Marketplace that want to participate for a total of 45 states.

Respondents per unit. See burden table.

Total Respondents. The total respondents were calculated by summing the product of the number of Units by the Respondents per Unit and total Spanish and Chinese Respondents.

Number of responses per respondent. Respondents will only be asked to respond once.

Hours per response. The revised Marketplace Survey is 83 items with an estimated completion time of 14.4 minutes (0.24 hours) (approximately 5.75 questions per minute). This estimate is based on the time it took to complete the Marketplace Survey in the psychometric test: 19 minutes for the phone administration and 14 minutes for the online administration for an average of 16.5 minutes to answer 95 questions. This is approximately 5.75 questions per minute.

Survey vendors. There will be a single survey vendor for all rounds of administration for the Marketplace Survey that is a CMS contractor. Thus, no vendor burden is associated with the Marketplace Survey.

Labor cost estimates. The Bureau of Labor Statistics reported the average hourly wage for civilian workers in the United States was $ 24.66 in November 2014[4]. See exhibit 4 for estimated burden costs.

Exhibit 4. Estimated burden hours and labor cost for Marketplace Survey beta test

Data Collection	Total Completes	Hours per Response	Total Burden Hours	Average Hourly Wage Rate	Total Cost Burden
SBMs (7)	8,400	0.24	2,016	$24.66	$49,714.56
SSMs (3)	3,600	0.24	864	$24.66	$21,306.24
SPM (7)	8,400	0.24	2,016	$24.66	$49,714.56
FFMs (27)	32,400	0.24	7,776	$24.66	$191,756.16
Total Beta Test	52,800	--	12,672	--	$312,491.52

---

[4] http://www.bls.gov/news.release/empsit.t19.htm