**TO:** Jennifer Park, Office of Management and Budget

**FROM:** Timothy Wojan **DATE:** 8/19/2014

**SUBJECT**: OMB CONTROL NUMBER: 0536-0071

 Non-substantive changes to Rural Establishment Innovation Survey (REIS) main study contact protocol based on assessment of completed survey yields.

 This memo requests a non-substantive change to the REIS contact protocol that will add an additional postcard reminder with web link as well as a total of 584 increased burden hours associated with this change. The anticipated completion rates given study budget restrictions are also discussed. Although the anticipated completion rate is lower than that assumed in the revised Supporting Statement, evidence from survey completions to date suggests that assumptions regarding the prevalence of rare events (e.g. application for a patent) were too conservative. Thus, there will be sufficient power to detect the effects (if they actually exist) of the types of rare events that are of interest to this study given the anticipated completions based on fielding results thus far. However, the lower response rates do increase concerns regarding nonresponse bias. Strategies to assess the seriousness of potential bias problems are discussed.

1. Change in contact protocol: A brief summary of REIS fielding results is reported in a table near the end of this memo. The AAPOR Response Rate 4 as of 7/31/14 was 18.3%. Since a large number of the web partial completes are fully usable other than the open ended comment section of the last question Response Rate 4 provides a realistic assessment of the response rate so far. Telephone interviewing began on 7/7/14 and has been most effective in directing respondents to the web to complete the survey. The completion rate of respondent directed to the web has averaged between 15% and 20%. Given the modest productivity of telephone attempts to generate completes to date, a second postcard reminder (Attachment E) with web link is suggesting as the most efficient way to boost the response rate. The attached graph of cumulative daily completes provides evidence of a temporary increase in the completion rate after the first postcard reminder was sent out. The timing for the mailing would correspond with a decline in productiveness of telephone contact anticipated to happen in the last two weeks of August. The additional burden associated with a second mailing is assumed to be 1 minutes per respondent mailed to the remaining noncontact sample of approximately 35,000 for a total of 584 hours.
2. Given the 18.3% as of 7/31/14 and anticipated productivity of phone and web completes, it is unlikely that the response rate of approximately 30% assumed in the Supporting Statement will be achieved. The most optimistic projection would be that 17.5% of the remaining sample of approximately 35,000 would generate an additional 6,125 full or partial completes bringing the total to roughly 15,000. A minimum required sample size of approximately 17,000 was anticipated in the Supporting Statement.

This minimum required sample size was estimated on the basis of testing for differences between rural and urban establishments with respect to rare events such as application for a patent. Examination of the data collected to date suggests that estimates of the rareness of these events—derived from European Union estimates of patenting in the services-producing sectors—was too conservative. We originally assumed that 3.183% of urban establishment would apply for at least one patent over the past 3 years and that 2.4574% of rural establishments would apply for a patent. The closeness of these two estimates was based on the observed higher patent application rate of manufacturing firms and the higher concentration of manufacturing firms in rural areas. In fact, the patent application rate by urban respondents observed so far is 8.1% and 5.62% by rural respondents. Given these rates the current sample size will provide a powerful test:

The POWER Procedure

Pearson Chi-square Test for Two Proportions

| **Fixed Scenario Elements** |
| --- |
| **Distribution** | Asymptotic normal |
| **Method** | Normal approximation |
| **Number of Sides** | 1 |
| **Null Proportion Difference** | 0 |
| **Group 1 Proportion** | 0.081 |
| **Group 2 Proportion** | 0.056 |
| **Group 1 Sample Size** | 2500 |
| **Group 2 Sample Size** | 5000 |
| **Alpha** | 0.05 |

| **Computed Power** |
| --- |
| **Power** |
| 0.991 |

Testing the more challenging comparison of the awarded patent rate will also be possible with an increase in total sample size to 12,000:given preliminary estimates:

The POWER Procedure

Pearson Chi-square Test for Two Proportions

| **Fixed Scenario Elements** |
| --- |
| **Distribution** | Asymptotic normal |
| **Method** | Normal approximation |
| **Number of Sides** | 1 |
| **Null Proportion Difference** | 0 |
| **Group 1 Proportion** | 0.045 |
| **Group 2 Proportion** | 0.035 |
| **Group 1 Sample Size** | 4000 |
| **Group 2 Sample Size** | 8000 |
| **Alpha** | 0.05 |

| **Computed Power** |
| --- |
| **Power** |
| 0.842 |

1. The lower than expected response rate so far does not appear to be a threat to the statistical power of this study. The more serious problem a lower response rate presents is the potential introduction of serious nonresponse bias. We will conduct nonresponse bias analysis using three different components of the original and collected sample:
	1. Nonresponse analysis of the full original sample will provide information on substantive differences in response rates across strata (metro/nonmetro, establishment size class, and industry membership at the NAICS two digit level). In addition , the original sample also contains information on establishment age and detailed 6-digit NAICS codes. Examining detailed industry characteristics (e.g., skill intensity, import penetration, patenting rates, innovation rates, etc.) will provide additional critical information on whether particular establishment characteristics associated with industry membership are associated with nonresponse.
	2. Nonresponse analysis of the proprietary SSI (Dunn and Bradstreet) sample. In addition to the analysis conducted for the full sample, the SSI dataset contains addition information on sales volume, establishment employment, company employment, and whether the establishment is part of a Fortune 1000 company.
	3. Analysis of refusal conversion survey data. Respondents who refused participating in the survey will have the opportunity to complete a brief 8 question survey. This instrument will provide explicit information on whether refusing establishments are substantially different than the responding sample.

|  |  |
| --- | --- |
| ***AAPOR Outcome Rate Calculator*** | **Rural Establishment Innovation Survey (REIS) Data Fielding Report**   |
| Last update 07-31-2014 | Final Disposition Codes | Category | Results |
| **Interview (Category 1)** |   |   |   |
| Phone completes | 1.1100 | I | 221 |
| Mail completes | 1.1200 | I | 3400 |
| Web completes | 1.1300 | I | 4218 |
| Phone partial completes | 1.2100 | P | 18 |
| Mail partial completes | 1.2200 | P | 0 |
| Web partial completes | 1.2300 | P | 1215 |
| **Eligible, non-interview (Category 2)** | 2.0000 |   |   |
| Refusal and breakoff | 2.1000 | RF | 261 |
| Refusal  | 2.1100 | RF | 2269 |
| Non-contact | 2.2000 | NC | 11175 |
| Respondent never available | 2.2100 | NC | 21 |
| Answering machine household-no message left | 2.2210 | NC | 1678 |
| Answering machine household-message left | 2.2220 | NC | 7659 |
| Deceased respondent | 2.3100 | O | 2 |
| Physically or mentally unable/incompetent | 2.3200 | O | 7 |
| Language problem | 2.3300 | O | 26 |
| **Unknown eligibility, non-interview (Category 3)** | 3.0000 |   |   |
| Always busy | 3.1200 | UH | 215 |
| No answer | 3.1300 | UH | 2911 |
| Call blocking | 3.1500 | UH | 102 |
| Not yet called | 3.2300 | UO | 15484 |
| **Not eligible (Category 4)** | 4.0000 |   |   |
| Fax/data line | 4.2000 | NW | 240 |
| Non-working number | 4.3100 | NW | 9 |
| Disconnected number | 4.3200 | NW | 699 |
| Temporarily out of service | 4.3300 | NW | 597 |
| Number changed | 4.4100 | NW | 317 |
| No eligible respondent | 4.7000 | IE | 420 |
| Other | 4.9100 | OT | 48 |
| Duplicates | 4.9200 | OT | 4 |
| **Total phone numbers used** |   |   | 53216 |
| I=Complete Interviews (1.1) |   |   | 7839 |
| P=Partial Interviews (1.2) |   |   | 1233 |
| R=Refusal and break off (2.1) |   |   | 2530 |
| NC=Non Contact (2.2) |   |   | 20533 |
| O=Other (2.0, 2.3) |   |   | 35 |
| e is the estimated proportion of cases of unknown eligibility that are eligible.  |   |   | 0.932 |
| UH=Unknown Household (3.1) |   |   | 3228 |
| UO=Unknown other (3.2-3.9) |   |   | 15484 |
|   |   |   |   |
| **Response Rate 1** | I / (I + P) + (R + NC + O) + (UH + UO) | 15.4% |
| **Response Rate 2** | (I + P) / (I + P) + (R + NC + O) + (UH + UO) | 17.8% |
| **Response Rate 3** | I / ((I + P) + (R + NC + O) + e(UH + UO)) | 15.8% |
| **Response Rate 4** | (I + P) / ((I + P) + (R + NC + O) + e(UH + UO)) | 18.3% |
|   |   |   |   |
| **Cooperation Rate 1** | I / (I + P) + R + O) | 67.4% |
| **Cooperation Rate 2** | (I + P) / ((I + P) + R + 0)) | 78.0% |
| **Cooperation Rate 3** | I / ((I + P) + R)) | 67.6% |
| **Cooperation Rate 4** | (I + P) / ((I + P) + R)) | 78.2% |
|   |   |   |   |
| **Refusal Rate 1** | R / ((I + P) + (R + NC + O) + UH + UO)) | 5.0% |
| **Refusal Rate 2** | R / ((I + P) + (R + NC + O) + e(UH + UO)) | 5.1% |
| **Refusal Rate 3** | R / ((I + P) + (R + NC + O)) | 7.9% |
|   |   |   |   |
| **Contact Rate 1** | (I + P) + R + O / (I + P) + R + O + NC + (UH + UO) | 22.9% |
| **Contact Rate 2** | (I + P) + R + O / (I + P) + R + O + NC + e(UH + UO) | 23.5% |
| **Contact Rate 3** | (I + P) + R + O / (I + P) + R + O + NC | 36.2% |

Cumulative completes by day and contact sequence