

**Longitudinal Study of
Unemployment Insurance
Recipients (LS-UI)**

OMB Supporting Statement

Part B: Collection of Information
Involving Statistical Methods

CONTENTS

PART B: COLLECTION OF INFORMATION INVOLVING STATISTICAL METHODS.....1

1. Respondent Universe and Sampling.....1

2. Analysis Methods and Degree of Accuracy.....4

3. Methods to Maximize Response Rates and Data Reliability.....6

4. Tests of Procedures and Methods.....12

5. Individuals Consulted on Statistical Aspects of the Design.....13

REFERENCES.....14

APPENDIX A: RESPONDENT MAILING MATERIALS

APPENDIX B: PRETEST MEMO

TABLES

B.1 Projected Precision for an Estimated Percentage of Recipients in Each Study Area for Each of the Three Rounds of Data Collection.....3

B.2 Projected Precision for an Estimated Percentage of Recipients in Each Study Area for Each Round for Selected Subpopulation Sizes.....3

B.3 Number of Completed Interviews Required for a Given Precision Level.....6

B.4 Estimates for Completion by Mode.....9

PART B: COLLECTION OF INFORMATION INVOLVING STATISTICAL METHODS

The U.S. unemployment insurance (UI) program aims to reduce financial hardships for unemployed workers, assist with reemployment, and ameliorate the negative effects of unemployment on the economy as a whole. The loss of a job poses major hardships for many workers and their families. Job losers often need to not only begin a potentially challenging search for new employment but also adjust their spending patterns and seek other sources of income. For qualified unemployed workers, UI benefits can help reduce the urgency for such adjustments. By providing temporary income support, UI benefits can smooth the transition to new circumstances, reduce financial distress, and provide workers with a buffer while they search for jobs. Furthermore, to reduce the potential incentive for UI recipients to prolong their unemployment, UI benefits are time-limited and provide only a partial replacement of lost earnings.

Understanding how workers adjust to changes in income during and after UI claim spells would enable policymakers to assess how well the program serves workers and refine it to meet the needs of unemployed workers while encouraging them to return to work. However, information about UI recipients is generally obtained from surveys that ask about experiences over a period of several years, which might not provide sufficient insight into the dynamic adjustments workers make after job loss or UI recipients' satisfaction with the program.

Given the importance of the UI program, the Office of the Assistant Secretary for Policy (OASP) of the U.S. Department of Labor (DOL) wants to understand the extent to which the program reduces recipients' financial hardships, the ways in which job search and reemployment expectations change during and after benefit collection, and customers' satisfaction levels with the program. As a first step, the Longitudinal Study of UI Recipients (LS-UI) will interview UI recipients in California to provide DOL with new insights about these issues. California has the largest number of unemployed workers of any state, and its UI system is a critical support for U.S. workers. For example, initial claims filed in California during the second week of June 2014 represented 20 percent of all initial claims filed that week nationwide.¹ While the statistics from the survey will not be representative of either California or the nation as a whole, the study will generate methodological insights that will be useful to DOL in the future should it desire to conduct potential surveys of UI recipients in other states.

To provide data about the experiences of UI recipients, OASP/DOL awarded a contract to Mathematica Policy Research to conduct the LS-UI. The study will address research questions in six broad topic areas: (1) adequacy of UI benefits, (2) reemployment expectations, (3) job search, (4)

total UI benefit usage, (5) employment outcomes, and (6) customer satisfaction.

Mathematica will conduct three surveys timed to coincide with the early, middle, and post-UI collection experiences of recipients in the Los Angeles metropolitan statistical area (MSA) and MSAs in the Central Valley area. The survey will follow a group of UI recipients for approximately one year to gain insight into the role that UI payments play in their lives. The LS-UI will gather information in short retrospective windows to reduce recall bias.

1. Respondent Universe and Sampling

This section describes the sampling design, potential respondent universe, sampling unit, estimated sample size, and expected response rate for the LS-UI.

a. Description of the Sampling Design

For the LS-UI, the sampling design will be stratified random samples of UI recipients in each of two purposively selected areas. The study will focus on UI claimants who received their first payment for UI benefits for a specific calendar week of unemployment that we refer to as a reference week. (The reference week is discussed more in Section B.2.) Within each study area, a sample of UI recipients will be selected in up to two cohorts (defined by a reference week) from a sampling frame based on administrative data extracts from the California Unemployment Benefit Services (CUBS) system obtained within a few weeks of the reference week. Ideally, only one cohort (based on a single reference week) will be needed to provide an adequate number of sample members for the study. However, if needed to ensure a sufficiently large sample for the study, samples will be selected from a second cohort, and the administrative data extracts for the sampling frames for the second cohort will be based on a reference week that is approximately six (6) weeks later than the reference week for the first cohort.

We have considered several factors in selecting the two areas for the LS-UI. First, we selected areas from California because of DOL's prior experiences receiving high-quality and timely data from the state. Within California, the Los Angeles MSA and Central Valley area were chosen because they are economically and geographically diverse, and they have large populations that can provide enough sample members for the study. The portions of the Central Valley for the study will be selected based on the likelihood that there will be enough UI claimants who receive their first payments for UI benefits to meet the requirements of the Office of Management and Budget (OMB) for minimum detectable differences. Because these two areas were selected purposively, the survey estimates will not be interpreted as generalizable to a broader population of recipients.

b. Description of the Potential Respondent Universe

The respondent universe is UI recipients in each of the two areas in California—Los Angeles and a collection of MSAs in the Central Valley—for specific reference weeks.

c. Sampling Unit

As described previously, the sampling unit is an eligible UI claimant who received a first payment for UI benefits for a specific week of unemployment. (It is possible that the state later makes a determination that the claimant was ineligible for that payment, but that would not be known at the time that the sample is drawn for the survey. Administrative data collected near the end of the study will be used to assess how prevalent this issue is, although it is expected to be uncommon.)

d. Population Frame

The sampling frame will be based on administrative data extracts for the reference week. The administrative data extracts will be requested from the California's UI office and pulled from the CUBS system.

e. Estimated Sample Size

The estimated sample size was determined based on the precision of an estimated outcome measure for the UI recipients in a reference week. The precision of an estimated outcome depends on the sample size and the value of the outcome. The nearer the prevalence of the outcome is to 50 percent, the larger the so-called margin of error (defined as the half-width of a 95 percent confidence interval) is around the estimate for a given sample size.

For each round of the survey, we need to account for the anticipated level of response in the projections for the precision of survey estimates. We have assumed an 80 percent response rate for the first round of data collection and an 85 percent response rate at the second round because of the recipient's cooperation in the first round. For the third round, we have assumed that 90 percent of the second round respondents will respond.

Based on these response rate assumptions, we have estimated that a sample size of 1,089 in each MSA for the first round of the survey will provide an adequate margin of error for estimates of outcomes—such as being reemployed at the time of an interview or use of reemployment services since the start of UI benefit collection—that occur with a 50 percent prevalence in the population plus or minus 3 percentage points for survey estimates in the first round (see Table B.1). For the second round, a projected sample size of 926 recipients (85 percent of 1,089 first round respondents) will achieve plus or minus 3 percentage points for estimates of outcomes that occur with a 33 percent (or 67 percent) prevalence. For the third round, a projected sample size of 833 recipients (90 percent of 926

Part B: Collection of Information Involving Statistical Methods

second round respondents) will achieve plus or minus 3.2 percentage points for estimates of outcomes that occur with a 33 percent (or 67 percent) prevalence. Outcomes that occur with a 50 percent prevalence will have a larger margin of error, whereas those that occur with a prevalence that is further from 50 percent (such as 25 or 75 percent) will have a smaller margin of error.

Table B.1. Projected Precision for an Estimated Percentage of Recipients in Each Study Area for Each of the Three Rounds of Data Collection

	Sample Size	Half-width of 95 Percent Confidence Interval for an Outcome Measured as a Percentage		
		25% / 75%	33% / 67%	50%
First Round	1,089	2.6	2.8	3.0
Second Round	926	2.8	3.0	3.2
Third Round	833	2.9	3.2	3.4

Source: Mathematica computations assuming a binomial distribution for the outcome measured. The equation for the confidence interval is $Z(\alpha) * [(P * (100 - P)) / n]^{0.5}$ where $Z(\alpha) = 1.96$, P is the percentage and n is the sample size for each round.

In Table B.2, we show the expected precision (as measured by the half-width of a 95 percent confidence interval for an estimated percentage) for subpopulations representing 25 percent, 33 percent and 50 percent of the study population based on completed interviews in each study area in each round.

Table B.2. Projected Precision for an Estimated Percentage of Recipients in Each Study Area for Each Round for Selected Subpopulation Sizes

	Sample Size	Half-width of 95 Percent Confidence Interval for an Outcome Measured as a Percentage		
		25% / 75%	33% / 67%	50%
First Round	1,089	2.6	2.8	3.0
Subpopulations				
50.0%	545	3.6	4.0	4.2
33.3%	363	4.5	4.8	5.1
25.0%	272	5.1	5.6	5.9
Second Round	926	2.8	3.0	3.2
Subpopulations				
50.0%	463	3.9	4.3	4.6
33.3%	309	4.8	5.3	5.6
25.0%	232	5.6	6.1	6.4
Third Round	833	2.9	3.2	3.4
Subpopulations				
50.0%	417	4.2	4.5	4.8
33.3%	278	5.1	5.5	5.9
25.0%	208	5.9	6.4	6.8

Source: Mathematica computations assuming a binomial distribution for the outcome measured. The equation for the confidence interval is $Z(\alpha) * [(P * (100 - P)) / k * n]^{0.5}$ where $Z(\alpha) = 1.96$, P is the percentage, n is the sample size and k is the subsample percentage.

f. Response Rates

The LS-UI is unique from many recent surveys conducted for DOL in terms of the age of the survey sample. While recent surveys conducted by Mathematica for DOL utilized sample that is generally two or more years old, the LS-UI will be selecting a sample that is comprised of UI recipients who are only weeks into their claim period. The quality of sample contact information is therefore, expected to be higher than usual. Also, the contractor will only attempt interviews with sample members at each round of data collection who have completed a prior round. Contact information will be updated at each interview. This strategy coupled with the expected high quality of sample contact information is expected to yield approximately 833 respondents with completed interviews at each survey round—a response rate of 80 percent. Mathematica achieved response rates above 80 percent in the Accelerated Benefits (AB) Demonstration conducted for the Social Security Administration which also had a sample of recent program enrollees. In that study, newly enrolled SSDI beneficiaries were sampled and response rates of 99 percent and 88 percent were obtained at baseline and at the 12-month follow-up, respectively.¹

2. Analysis Methods and Degree of Accuracy

a. Description of Stratification

Within each area, the sampling frame will consist of an administrative data extract of UI recipients for up to two reference weeks (separated in time by six calendar weeks). Because the sample will be drawn from up to two weeks only, chance events such as a mass layoff might influence the sample. The analysis will explicitly state that the survey estimates relate to UI claimants who received their first payments for UI benefits in the Los Angeles or Central Valley areas for those weeks, but the findings cannot be generalized to the typical set of UI claimants who receive their first payments for benefits over a broader period of time in those areas. We will additionally look for evidence of a mass layoff in the sample by identifying whether UI recipients are particularly likely to report a separating employer so that we can understand the composition of the sample and, if needed, ensure diversity among survey sample members through the sampling process. Within each reference week, the sampling frame will be implicitly stratified by a few characteristics, such as variables related to age, the Worker Profiling and Reemployment Services score and/or, if needed, the pre-UI employer of the UI recipient. Implicit stratification is a process in which the sampling frame is sorted two or more implicit stratification factors and a sequential selection procedure (similar to systematic sampling in which every n^{th} unit is selected) is used to select the sample. The implicit stratification will result in an approximate proportional allocation of the sample across the implicit stratification factors without forming explicit

¹ Participants in the AB Demonstration were randomly assigned to three study groups. Two study groups were eligible to enroll in health insurance plans.

strata and specifying sample size for these explicit strata, which can result in unequal selection probabilities and increases in the sampling variances. This implicit stratification will result in having the sample mirror the distributional characteristics of the sampling frame for the implicit stratification factors (enhancing the face validity of the study samples) and will preserve the equal selection probabilities within the reference week samples.

b. Description of Sample Selection Methodology

The sample in each reference week will be selected using a sequential selection procedure that permits equal probability sampling with implicit stratification and permits unbiased estimation of the sampling variance for survey estimates. Dr. James Chromy developed this method (Chromy 1979), and its implementation is available in the SAS software package.

c. Procedure for Variance Estimation

The sample design is a stratified random sample and the computation of survey estimates based on the sample requires the use of survey data analysis procedures. These survey data analysis procedures are available in the SAS, STATA, and SUDAAN statistical software packages. The sampling variance will be computed for nonlinear estimates (such as proportions, percentages, means, and regression coefficients) using the Taylor series linearized expansion of the survey estimator and the explicit equations for stratified random sampling. The data file of respondents will include the stratification parameters to permit the computation of correct sampling variances.

d. Description of Plans for Analyses, Including Key Variables and Proposed Statistical Tests

The analysis of LS-UI data will use multivariate techniques to examine the extent to which demographic variables and economic factors faced before an unemployment spell relate to UI recipients' experiences of financial hardship. The analysis will also examine factors that relate to reemployment and earnings over time. Consistent with the economics literature studying earnings, the analysis may apply a transformation to earnings when it is the dependent variable to reduce the influence of extreme values of income in our estimates (see, for example, Pence [2006] and Heckman et al. [2003]). As with the tabulations discussed in Section A.16, all models will be estimated separately for each study area and include weights so the estimates are representative to UI recipients in each area. The samples will not be pooled because the resulting statistics would not be relevant to policymakers. Estimating models separately will also allow the relationships between variables to be calculated based on UI recipients in each area. Thus, the models are likely to better fit the data through separate estimation. We will compare the results of the models to describe the experiences of UI recipients in the two contexts. However, there is a possibility that certain

types of UI recipients, such as those who move frequently or do not have a computer, will be less likely to respond to the survey or have missing administrative data for reasons that are not accounted for in the weights. Both of these sources of missing data would cause nonresponse bias. The discussion of the analysis will include appropriate caveats to indicate that there is a potential for nonresponse bias.

For binary outcomes such as whether a UI recipient has experienced financial hardship (constructed from responses to survey data items), a probit or logit model specification will be estimated of the form:

$$(1) \quad P(\text{Experienced financial hardship}) = F(Z\beta),$$

where an indicator variable for experiencing financial hardship is regressed on a vector of variables (Z) that influence the outcome, β is a vector of coefficients to be estimated, and F is a probability density function (such as the normal distribution for probit analysis or the logistic distribution for logit analysis). Z -tests for the estimated coefficients will be used to determine whether factors are statistically significantly correlated with the probability of experiencing financial hardship.

For continuous outcomes such as earnings or transformed earnings at each round, linear regression models will be estimated of the form:

$$(2) \quad Y = Z\beta + \varepsilon,$$

where the outcome Y is regressed on a vector of covariates (Z) and ε is a random error term. T -tests for the estimated coefficients will be used to determine statistical significance.

Variables that may be included in Z as covariates for the models in Equations (1) and (2) are survey mode, demographic characteristics, economic characteristics, and some pre-UI job characteristics that are common in the literature on earnings and employment (see, for example, Addison and Blackburn [2000] and Ehrenberg and Oaxaca [1976]). We will assess whether there are differences in nonresponse rates and means of key measures by survey mode. If there are differences, then we will include an indicator for the survey mode as a covariate. Demographic and economic characteristics include sex, age, race/ethnicity, education level, family size, marital status, employment status of spouse if married or partner if living with an unmarried partner, and other household income. Pre-UI labor market characteristics include industry, occupation; weekly earnings at the pre-UI job; job tenure; and previous receipt of UI. Measures of local economic conditions at the MSA or county level before the unemployment spell, such as the unemployment rate or population size, are common covariates in the literature but will be excluded from the analysis. There will not be sufficient geographic variation to control for these local covariates in the model because the Los Angeles MSA will be analyzed separately from the Central

Valley area. Also, because the sample intake for the survey will be completed in a very narrow time window, the UI recipients within an area are likely to face similar economic conditions.

e. Minimal Substantively Significant Effect Sizes

Table B.3 shows the 95 percent confidence interval around an outcome occurring for some percentage of the population (such as benefit exhaustion or being satisfied with the program). For example, with a sample size of 288, the “margin of error” for an outcome of 25 percent is ±5 percentage points. As described previously, for some outcomes (such as being reemployed at the time of an interview or use of reemployment services since the start of UI benefit collection), the projected percentage in the study population is about 33 percent. The target number of completed interviews is 833 for each labor market area after the third round (that is having responded in all three rounds), which will achieve a substantively significant level of accuracy of approximately ±3.2 percentage points for a 95 percent confidence interval.

Table B.3. Number of Completed Interviews Required for a Given Precision Level

95 Percent Confidence Interval	Outcome Measured as a Percentage		
	25 / 75%	33 / 67%	50%
±5.0 percentage points	288	341	384
±3.2 percentage points	703	833	938
±3.0 percentage points	787	933	1,050
±2.8 percentage points	919	1,089	1,225

Source: Mathematica computations assuming a binomial distribution for the outcome measured. The equation for the confidence interval is $Z(\alpha) * [(P * (100 - P)) / n]^{0.5}$ where $Z(\alpha) = 1.96$, P is the percentage and n is the sample size.

f. Unusual Problems Requiring Specialized Sampling Procedures

This study does not require specialized sampling procedures.

3. Methods to Maximize Response Rates and Data Reliability

The methods we will use to maximize response to each round of survey data collection and ensure the reliability of the data collected are discussed in this section.

a. Response Rates

As discussed in Section B.1, sample for the LS-UI will be selected from the Los Angeles MSA and the Central Valley area. DOL and Mathematica have carefully reviewed California’s UI program characteristics to determine its suitability for the study. DOL made initial contacts with California to secure interest in and commitment to participating in the study, and Mathematica will follow up these calls to specify and clarify administrative data content

and timing needs. The request for state participation and the release of administrative data is also supported by DOL's UIPL 23-12, which requires states to disclose unemployment compensation (UC) information, including confidential claim information, needed for Office of Management and Budget (OMB) approved evaluations of UC programs conducted by DOL. The regulation allows Mathematica, acting as DOL's agent, to obtain the UI claims data needed to conduct the study. Resources have been included in the evaluation contract to reimburse states for their reasonable costs associated with disclosures of data for this evaluation.

As with any survey, some nonresponse among sample members will occur. DOL and Mathematica will take steps to maximize response rates and to address potential bias through nonresponse analysis. The study will maximize participation among sample members by adopting practices that have been successfully used in studies of similar populations. The methods employed will address all types of individual nonresponse, including failure to locate the individual and sample members' refusals to participate.

Contact with sample members. Using an envelope displaying the DOL logo, Mathematica will mail an advance letter printed on DOL letterhead and signed by a senior DOL official to sample members before attempting to contact them by telephone. The official logo and letterhead will help to capture the attention of the sample member and establish the legitimacy of the survey. Mathematica's address will be provided as the return address on the envelope to help quickly process returned mail and start any necessary locating procedures. The advance letter will (1) introduce the study and emphasize its purpose (to understand the experiences of persons who become unemployed and who receive UI benefits), (2) highlight DOL as the study sponsor, (3) explain the voluntary and private nature of participation, (4) extend the incentive offer, (5) provide web survey log-in information, and (6) give a toll-free number for those without web access or who prefer to complete the survey by telephone. An information sheet providing answers to questions that sample members have about the study will be included as part of the initial mailing. The advance letter will be followed up with timed reminders to nonrespondents offering the option to complete the survey via the web or by telephone. Each of these materials will emphasize the differential incentive to encourage respondents to complete the survey by web or to call in to complete it. Copies of the advance letter, information sheet, and reminder materials (postcard and letter) that will be sent to sample members are provided as Appendix A.

Before this mailing, staff in Mathematica's Survey Operations Center (SOC), including interviewers, project supervisors, monitors, and locators, will receive comprehensive training on how to address respondents' questions about the study and how to administer the questionnaire. The project team will develop a list of frequently asked questions (FAQs) for telephone interviewers who respond to questions from sample members. These FAQs will be included in the operational procedures manual developed for

computer-assisted telephone interviewing (CATI) interviewer training, and will be integrated into the CATI instrument. Interviewers will be able to access the FAQs at any time during the CATI survey. A version of the FAQs will also be available on the log-in screen for web survey respondents to access throughout the survey.

Locating sample members. A key component to obtaining a high response rate is locating sample members. Before mailing the advance letter, we will use an independent vendor to verify the contact information we received for the sample from the administrative data extracts. Specifically, we will check the administrative data against current address databases such as Accurant. This step will assist us in quickly determining cases that will require locating efforts. Additionally, providing the Mathematica return address on the advance letter envelopes will produce returned mail, sending the cases at those addresses directly into the locating process. For cases requiring locating, we will use procedures that have proven successful in other Mathematica studies. These efforts include searching independent databases and, if needed, checking with neighbors and family members to locate respondents. To maintain the privacy of sample members, when talking with contacts the specific purpose of the call will not be disclosed. It will be stated that the attempt to reach the sample member is for an important study sponsored by the government.

To assist in gathering updated contact information for the planned second and third interviews, as part of each survey, we will collect detailed contact information for the sample member and up to three friends and/or relatives who will know how to contact them. The contractor will also request sample members' cell phone numbers and email addresses, and ask permission to contact them (privately and securely) if we have difficulty reaching them.

Gaining and maintaining cooperation. A key component to achieving high response rates is gaining cooperation after locating respondents. Mathematica's interviewers are highly trained in establishing rapport with gatekeepers, gaining cooperation, and averting refusals. Sample members who are difficult to contact and who have not yet completed the survey on the web will receive a reminder postcard one week after the advance letter. This reminder postcard will provide the sample member with a call-in number to complete the survey on the telephone or receive their web survey log-in information. One week later, a reminder letter will be sent to sample members who have not responded. Sample members who refuse to participate will be sent a targeted refusal conversion letter and email designed to address their specific concerns. If they still do not complete the survey following these outreach efforts, a trained refusal conversion interviewer will attempt to contact the sample member and gain his or her cooperation. Similar to the advance letter, both the reminder letter and refusal letters will be sent on DOL letterhead.

Multi-language survey administration. In anticipation of multi-language needs, all instruments will be translated into Spanish and bilingual interviewers will be trained to conduct the CATI interview in Spanish. Mathematica will evaluate the need for translators for languages other than Spanish and will respond to these on a case-by-case basis, as determined in consultation with the DOL project officer. The costs of using outside translation or interpreting services will be considered in these determinations. Mathematica employs staff fluent in a wide range of languages and who have experience conducting interviews in many languages.

Incentives for survey participants. To maximize response rates and maintain data quality while controlling survey operations costs, an incentive will be offered at each round of data collection. A combination of pre- and post-paid incentives is proposed for the LS-UI. At round one, a \$5 bill will be included in the advance mailing that invites sample members to participate in the study. All LS-UI sample members will be sent the \$5 pre-payment with their advance letters and then, depending on the mode of completion, will receive either an additional \$25 for completion on the web or for calling in to complete, or an additional \$15 for interviews initiated by an interviewer. This two-tiered incentive offer encourages respondents to complete using the less expensive survey modes. We anticipate that due to the higher incentive and the high degree of comfort much of the general population has with the Internet, a substantial number of respondents will opt to complete the survey in this mode. Log-in information for the web survey will be provided to sample members in their advance letters, at which time the survey will go live on the web.

Since the study will only follow first interview completers at subsequent rounds of data collection, rounds two and three will use a post-pay incentive only. The rationale for this is that the researchers will have established a relationship and trust with sample members and the pre-payment will not be needed.

Incentive payments have been found to contain evaluation costs by significantly reducing the number of calls required to resolve a case. Studies offering incentives show decreased refusal rates and increased contact and cooperation rates. Jäckle and Lynn (2007) found that incentives increased the participation of sample members more likely to be unemployed. Singer et al. (2000), also support the use of incentives to achieve high response rates by increasing the propensity of sample members to respond. This increased propensity to respond was demonstrated in an experiment conducted by Mathematica for DOL's National Evaluation of the Trade Adjustment Assistance Program (OMB Control Number 1205-0460) in 2008. In that experiment different levels of incentives were offered to sample members who were not responding to survey outreach attempts. Nonrespondents were randomly assigned to three groups: (1) a group that was offered an incentive of \$25, the same amount as paid to respondents;

(2) a group that was offered an incentive of \$50; and (3) a group that was offered an incentive of \$75. The experiment found that the response rate was 9.4 percentage points higher with an incentive of \$50 than an incentive of \$25, a difference that was statistically significant; the response rate was 15.0 percentage points higher with an incentive of \$75 than an incentive of \$25.

Budgetary constraints preclude the contractor from offering more than \$30 for web and call-in responders and \$20 for LS-UI respondents who complete the survey after interviewer outreach. However, in an incentive experiment from the 1996 panel of the Survey of Income and Program Participation suggests that the \$20 level meets the minimum level required to improve response among adults. That experiment showed that a \$20 incentive significantly increased response rates, while a \$10 incentive had no effect relative to those who received no incentive. Burghardt and Homrighausen (2002) found response rates for the third follow-up survey of youth in the National Job Corps Study were low with only a \$10 incentive, but increased after OMB approved an increase in the incentive to \$25. The cost per completed interview was nearly 20 percent lower compared to interviews conducted at the lower incentive level. Part A of this clearance package provides additional justification for the incentive payment.

For the LS-UI, it is estimated that 60 percent of the completed first-round surveys will be completed via the web, reaching 65 percent and 70 percent in the subsequent rounds. These estimates are presented in Table B.4. In some surveys we have conducted, more than 90 percent of the surveys have been completed using the web. For the LS-UI, we have assumed a lower proportion because the population is diverse by age and has lower-than-average incomes—both elements correlated with reduced web usage.

Table B.4. Estimates for Completion by Mode

Field Period	Web	Interviewer-Initiated CATI
1st Round	60%	40%
2nd Round	65%	35%
3rd Round	70%	30%

All mailing materials will prominently mention the incentive in order to capitalize on all of its benefits. Interviewers will also emphasize the incentive as they attempt to gain the sample members' cooperation.

Survey length. The LS-UI survey questionnaire will be designed to be easy to complete, using questions written in clear and straightforward language. Survey questions will be salient and minimize burden by not asking about information across multiple surveys that would not have changed (such as race) and prefilling information learned in the first interview into follow-up interviews. Burden is also reduced by using

questionnaire logic and skip patterns that direct respondents to relevant questions and bypass irrelevant ones. The estimated average time required for the respondent to complete the survey, either on the web or by telephone, is 25 minutes.

Interviewer training. Mathematica has a large number of experienced survey operations staff who have worked on previous studies conducted for DOL (as interviewers, supervisors, and monitors). These staff are familiar with similar questionnaire content and are sensitive to the difficulties faced by job seekers and unemployed individuals. To the extent possible, Mathematica will assign these experienced staff to the study. In addition to standard general interviewer training, all interviewing staff will participate in extensive project-specific training. This training specific to LS-UI will include a review of the project background, frequently asked questions, standards for gaining and maintaining cooperation with sample members, a thorough review of the questionnaire, multiple role-playing scenarios to practice survey administration, and refusal aversion. This training will also focus on the importance of being sensitive to sample members' concerns and situations, while still remaining impartial. Interviewers will not be permitted to work on the study until they have been certified as prepared through supervised paired-practice sessions.

b. Nonresponse Bias Analyses

We will conduct a nonresponse bias analysis to provide some indication of whether the potential for nonresponse bias exists, an indication of the individual data items and specific populations for which survey estimates might have a greater potential for bias, and the possible extent of the potential for nonresponse bias in survey estimates. However, because survey data will not be available for nonrespondents, we cannot be certain if bias does or does not exist in the survey estimates.

For the nonresponse bias analysis, we will use the various data collected in this study to compare the characteristics of respondents and nonrespondents using administrative data as an assessment of the potential for nonresponse bias. Administrative data (including demographic and employment history information available in the administrative records) will be available for all sampling frame members and will be the most useful data to define the subgroups for the nonresponse analysis.

For the nonresponse bias analysis, we plan the following steps:

1. Compute response rates for key subgroups of UI recipients based on the administrative data.
2. Compare the weighted distributions of respondents and nonrespondents (for administrative data characteristics using the unadjusted base weight).

3. Identify the characteristics that best help predict nonresponse through a chi-square automatic interaction detection (CHAID) analysis and logistic regression modeling.
4. Use this information to generate nonresponse weight adjustments.
5. Compare the distributions of respondents using the fully response-adjusted analysis weights for administrative data characteristics with the distributions for the comparable full sample weights using the unadjusted sampling weights.

Mathematica will summarize the results of the nonresponse bias analysis and will submit a memo describing the analysis procedures used and an assessment of the potential for nonresponse bias. The final report will include a brief appendix of these findings.

c. Plans for Nonresponse Adjusted Weights

Logistic regression modeling is commonly used to develop adjustment factors for nonresponse, also known as response propensity modeling. Response propensity modeling using logistic regression can be viewed as an extension of the classical weighting-class nonresponse adjustment procedure that makes it possible to include more factors (that is, binary, categorical, and continuous factors) in nonresponse adjustments. To simplify the process, CHAID is commonly used to assist in identifying potentially significant interactions among the subgroups or factors available for all individuals. We plan to use CHAID, with the initial sampling weights, to help identify the interactions.

The CHAID algorithm partitions the sample in a hierarchical fashion, with each successive splitting of the sample identified by CHAID. CHAID uses the chi-square statistic with the proportion responding defined as the dependent variable to determine the partitioning of the sample with the largest value for the statistic among all possible partitions by the factors available. After the initial partitioning, the chi-square statistic is again used to identify additional partitions subject to predetermined restrictions (for example, a minimum partition size).

Next, we develop variables that reflect the interaction terms identified through the CHAID analyses, and use these variables in forward and backward stepwise logistic regressions to eliminate redundant interaction variables and to identify the most-significant interactions. The stepwise logistic regressions are conducted using SAS software with normalized weights. However, the SAS software for stepwise logistic regression does not account for the sampling design. Hence, we will use the survey data analysis procedures in SAS or SUDAAN to develop the final model, so variance estimates for the coefficients reflect the sampling design. Goodness-of-fit for the final model is assessed using the percentage of concordance and

discordance, the R-square for the model, and the Hosmer-Lemeshow goodness-of-fit test statistic.

The final response propensity model described earlier will be used to identify factors associated with nonresponse and to compute the appropriate nonresponse adjustment factors for the sampling weights. The inverse of the predicted propensity to respond will be used as an adjustment factor to the initial sampling weights. These response-adjusted weights will then be post-stratified to totals computed using the full sampling frame and will be the final analysis weights.

d. Procedures to Handle Item Nonresponse

Because of the descriptive nature of the study, no imputations are planned for missing data. The tabulations will show the number of respondents with data.

e. Reliability of Data Collection

The LS-UI survey includes questions that have been tested and successfully used in the field by other recent studies, such as the Trade Adjustment Assistance Study Follow-Up Survey (OMB number 1205-0460), the Individual Training Account 2 Follow-up Questionnaire (OMB 1205-0441), and the COBRA Subsidy Study (OMB 1219-0001). In addition, it contains questions that have a long history of successful administration on the National Longitudinal Survey of Youth. Development of the LS-UI survey questionnaire benefited from reviews by staff at DOL, Mathematica, and the Center for Human Resource Research at the Ohio State University, and members of the project's technical working group. Furthermore, it was comprehensively pretested with a nonstudy sample of UI recipients. A memo detailing pretest results is included as Appendix B.

The use of CATI and web platforms will also increase reliability and data completeness. In both cases, certain questions will require an answer before the respondent can proceed. Along with required responses to critical questions, programmed logic such as probes, verifications, and consistency checks in both modes will further improve data reliability. The only difference between the two modes will be some adjustment in the text to accommodate self-administration versus the wording used by a telephone interviewer.

Certain information—such as the UI claim date, job separation date, and employer name—will be prefilled at relevant questions and verified with the respondent to help him or her focus consistently on the selected claim and job of interest.

Finally, interviewing supervisors will monitor at least 10 percent of each interviewer's work using silent call-monitoring equipment and video monitors that display the interviewer's screen. Interviewers' performance will be evaluated based on this monitoring and any performance issues that arise

will be discussed with a supervisor. Retraining and/or reassignments will be provided as needed.

f. Justification for Using a Sample Rather than Systematically Collecting Data for the Entire Respondent Population

The proposed sample size and sampling procedure will result in sufficient data for the planned analysis. The study cannot justify the respondent burden and the costs and time required for systematically collecting data for the entire respondent population.

4. Tests of Procedures and Methods

Before fielding, questionnaires for LS-UI will be pretested to evaluate the clarity of the questions asked, identify possible modifications to either question wording or question order that could improve the quality of the data, estimate respondent burden, and assess the overall data collection process. To best test the questionnaire, respondents who closely mirror the proposed sample of UI recipients will be recruited. Because a sample selected from UI administrative data files will not be available due to the timing, Mathematica will solicit staff for referrals of friends and family members who are receiving or recently received UI benefits. This approach has been used in the past with great success. Ensuring that pretest respondents are similar to the targeted sample will enable researchers to anticipate and properly address any issues sample members face during survey administration. Pretest respondents will be assured of the same level of privacy as full study participants.

We will pretest the questionnaire on up to nine UI recipients. To the extent possible, we will try to recruit and interview a mix of pretest respondents who are in the early, middle- and post-claim periods of UI receipt to mirror as closely as possible our survey sample respondents. This will enable us to test all questions that will be asked across the three rounds of data collection.

Pretest interviews will be monitored and recorded to identify questions that were problematic for interviewers or respondents and, at the conclusion of each interview, interviewers will debrief the respondent to gain additional insights. Mathematica's survey director will also conduct a debriefing session with interviewers upon completion of all pretests to obtain their perspectives on how well the survey instrument worked and where improvements are needed. This intensive approach will enable us to assess the effectiveness of the instrument and make any needed adjustments. This kind of debriefing has proven to be invaluable to similar data collection efforts.

The pretests will be conducted by telephone using hard-copy instruments due to the time required to program CATI and web surveys. When programmed, rigorous tests of the CATI and web applications will be

conducted. Pretest sample members will receive a \$30 post-paid incentive payment for completion. We will try to assess the pretest respondents' reactions to receiving a pre-paid cash incentive, but will not be able to implement the pre-pay/post-pay strategy planned for the main data collection due to time restrictions.

5. Individuals Consulted on Statistical Aspects of the Design

To ensure that the best decisions were made regarding the statistical aspects of the design, project staff from the Mathematica Policy Research evaluation team, as well as members of a Technical Working Group (TWG) contributed to the sampling design. The experts consulted are listed below, along with telephone contact information. All consultations were paid under the study's contract.

Mathematica Policy Research Evaluation Team Staff		
Ms. Julita Milliner-Waddell	Project /Survey Director	(609) 275-2206
Dr. Karen Needels	Senior Researcher	(541) 753-0201
Dr. Frank Potter	Senior Fellow	(239) 558-5956
Dr. Walter Nicholson	Senior Fellow	(239) 774-3693
Dr. Joanne Le	Researcher	(510) 830-3727
Dr. Stephen Wandner	Visiting Scholar at the Urban Institute and W. E. Upjohn Institute for Employment Research	(301) 785-6670
Dr. Randall J. Olsen	Center for Human Resource Research at the Ohio State University (CHRR)	(614) 442-7348
Dr. Randall J. Olsen	Center for Human Resource Research at the Ohio State University (CHRR)	(614) 442-7348
TWG Members		
Dr. Rich Hobbie	National Association of State Workforce Agencies	(202) 434-8020
Dr. Christopher O'Leary	W. E. Upjohn Institute for Employment Research	(269) 385-0407
Polly Phipps	Bureau of Labor Statistics (BLS), U.S. Department of Labor	(202) 691-7513

REFERENCES

- Addison, J. T., and Blackburn, M. L. (2000). The effects of unemployment insurance on postunemployment earnings. *Labour Economics*, 7(1), 21-53.
- Chromy, J. R. "Sequential Sample Selection Methods." *Proceedings of the American Statistical Association, Survey Research Methods Section*, 1979, pp. 401-406.
- Ehrenberg, R. G. and Oaxaca, R. L. (1976). Unemployment insurance, duration of unemployment, and subsequent wage gain [Electronic version]. *American Economic Review*, 66(5), 754-766.
- Heckman, James J., Lance J. Lochner, and Petra E. Todd (2003). Fifty years of Mincer earnings regressions. National Bureau of Research: No. w9732.
- Jäckle, Annette, and Peter Lynn. "Respondent Incentives in a Multi-Mode Panel Survey: Cumulative Effects on Nonresponse and Bias." Working paper presented to the Institute for Social and Economic Research, University of Essex, Colchester, United Kingdom, 2007.
- Pence, Karen M. (2006). "The role of wealth transformations: an application to estimating the effect of tax incentives on saving," *Contributions to Economic Analysis & Policy*: Vol. 5(1).
- Singer, Eleanor, John Van Hoewyk, and Mary P. Maher. "Experiments with Incentives in Telephone Surveys." *Public Opinion Quarterly*, vol. 64, no. 2, summer 2000, pp. 171-188.