B. PROJECT GATE II

COLLECTION OF INFORMATION INVOLVING STATISTICAL METHODS

1. Respondent Universe and Sampling

For each of the three random assignment sites, the study population includes all individuals who applied for GATE II services after March 2009. We expect the size of the universe to be 1,980 individuals based on current projections of GATE II enrollment as of the end of random assignment in mid-2011. In Alabama and Virginia, half of the eligible applicants will be randomly assigned to the group that are offered GATE II microenterprise services (the program group) and half will be assigned to the group not offered GATE II services but who can seek pre-existing microenterprise services (the control group). Random assignment is conducted at IMPAQ using a computer program that uses a random number generator with the system clock time in milliseconds as the random number seed. When the states notify IMPAQ that a particular GATE-eligible individual is ready for random assignment, IMPAQ performs the random assignment and notifies the state whether the individual was randomly assigned to the Treatment group or the Control group. In North Carolina, three-quarters of the eligible applicants will be randomly assigned to the program group, and one-quarter will be randomly assigned to the control group. Table 1 below provides estimates of the number of individuals who will be enrolled in each site and overall.

TABLE 1
EXPECTED NUMBER OF ELIGIBLE APPLICANTS BY SITE

Site	Expected Number of Applicants
Alabama	70
North Carolina	1350
Virginia	560
Total	1,980

Follow-up surveys at approximately 18 months following random assignment will be attempted with all 1,980 sample members. Based on experiences with similar surveys, we expect to obtain at least an 80 percent response rate. For example, in the original Project GATE evaluation conducted for DOL/ETA by IMPAQ, the response rate was 82% for the first of three waves of follow-up survey (at 6 months after random assignment), 88% for the second wave (at

18 months after random assignment), and 81% for the third wave (at 60 months after random assignment).[1] When implementing the second wave of the Project GATE survey, we attempted to contact only the 82% of the participants who completed the first wave. Thus, the net response rate after two waves was 88% of 82% of the original sample, or 72% of the original sample. In the current GATE II evaluation, we have a single wave of follow-up survey at approximately 18 months after random assignment. We will attempt to contact the entire original participant sample, not just the 82% or so who might have completed a 6-month survey if we had implemented one in GATE II. Thus, we expect our response rate at 18 months in the GATE II evaluation to be greater than 72% of the original participant population. Since we will be contacting the additional 18% (the difference between 82% and 100%) of GATE II participants when we conduct the GATE II follow-up survey, we only need to succeed in obtaining completed surveys from 8% of that 18% (a "subsample" response rate of 8/18 or 44%) in order to bring the 72% completion rate up to 80%.

2. Procedures for the Collection of Information

a. Statistical Methodology, Estimation, and Degree of Accuracy

The primary objective of the GATE II evaluation is to provide statistically valid and reliable estimates of the incremental effects of GATE II self-employment services on key outcomes, including self-employment, weeks worked, and earnings. Use of a classical experimental design, in which applicants are assigned randomly to program and control groups, will ensure that measured impacts represent valid estimates of the effects of the demonstration services. The measured impacts will be internally valid for the three sites that implemented a random assignment design (North Carolina, Virginia, and Alabama)[2]. Since the GATE II states have been chosen purposively, combined with the fact that the study has small sample sizes, GATE II impacts cannot be generalized to a wider population with a known degree of statistical precision.

Impacts will be estimated by computing differences in mean outcomes between individuals in the program group (that is, those offered GATE II services) and individuals in the control

---

[1] Project GATE Final Report: http://wdr.doleta.gov/research/FullText_Documents/Growing%20America%20Through%20Entrepreneurship%20-%20Final%20Evaluation%20of%20Project%20GATE.pdf

[2] As previously described, Minnesota did not implement a random assignment design; every applicant who was eligible for GATE II was offered GATE II services. A quasi-experimental evaluation would have been possible, but lack of consent from comparison group members to participate in the evaluation study precluded obtaining adequate data for impact analysis.

group. Although differences in distributions of outcomes between program and control group members provide the most complete understanding of impacts, differences in mean outcomes provides a widely accepted statistic that summarizes the distribution. Simple differences of means will be computed, but we will also adjust for random differences at intake using multivariate regression, as described in Section 16a of this document. The regression adjustments will increase the precision of the impact estimates. More detail on estimation procedures is included in our discussion of tabulation plans in Part A, under item 16a (pages 17-21).

Given this design the main question is whether the impact estimates will be precise enough to detect policy relevant impacts. Using sample sizes projected to be available based on tracking enrollment in GATE II at each of the project sites, Table 2 presents minimum detectable effects for comparisons across the full sample (i.e., 1327 program group members and 653 control group members) for four key labor market outcomes: (1) the percentage ever self-employed, (2) the percentage ever employed in either wage and salary employment or self-employment, (3) the amount of UI benefits received, and (4) total earnings from both self-employment and wage and salary employment over a 12 month period.

TABLE 2
MINIMUM DETECTABLE EFFECTS FOR KEY LABOR MARKET OUTCOMES AT 18 MONTHS AFTER RANDOM ASSIGNMENT

	Alabama	North Carolina	Virginia	Combined
Projected Sample Size	70	1,350	560	1,980
Likelihood of Self-Employment	33.1	8.7	11.7	6.6
Likelihood of Employment	26.6	7.0	9.4	5.3
UI Benefit Amounts Received	$1,331	$350	$471	$266
Total Earnings	$9,982	$2,625	$3,529	$1,997

NOTE: The calculations assume (1) a 95 percent confidence level with an 80 percent level of power; (2) a one-tail test; (3) 45 percent of the control group will become self-employed; (4) 80 percent of the control group will be employed in either wage and salary employment or self-employment; (5) a standard deviation of $2,000 for amount of UI benefits received; (6) a standard deviation of $15,000 for total earnings from salary employment or self-employment; and (7) a survey response rate of 80 percent. These assumptions are based on analyses from two similar studies of self-employment programs: 1) A New Reemployment Strategy, Final Report on the UI Self-Employment Demonstration (1995), Unemployment Insurance Occasional Paper No. 95-4. Washington, DC: U.S. Department of Labor; and 2) Growing America through Entrepreneurship: Findings

from the Evaluation of Project GATE (2008), Employment and Training Administration Occasional Paper 2008-08.

The minimum detectable effects (MDE) are calculated using the following formula: $$MDI = \alpha\sigma\sqrt{\frac{2(1-R^2)}{rn}}$$ where $\alpha = 2.5$ for a one-tail test, σ is the standard deviation of the variable, R^2 is the variance explained by the regression model, r is the response rate, and n is the size of the program and control group (Bloom, 1995).

Based on a review of results from previous studies of self-employment program impacts, we believe that a sample of 1327 program and 653 control group members will meet precision targets for impacts for the full sample and key subgroups. As seen in the table, the minimum detectable impact is 6.6 percentage points for the percentage ever self-employed, 5.3 percentage points for the likelihood of being employed in either wage and salary employment or self-employment, $266 for the amount of UI benefits received, and $1,997 for total earnings.

As noted, in addition to estimating the overall impact of GATE II, it is important to evaluate the impact the program for its two key subpopulations (rural and older workers) and, if feasible, assess if the program had differential impacts by other key characteristics (race, education, prior-self-employment experience, etc). Subgroup analyses will help assess whether certain participant subgroups are more likely to benefit from program participation than others or whether the program is equally effective across all participant subgroups. To help assess whether it is feasible to detect statistically significant and meaningful program impacts for key subgroups, Tables 3 and 4 present the MDE for key labor market outcomes by rural/older participant status and by key participant characteristics, respectively.

TABLE 3
MINIMUM DETECTABLE EFFECTS FOR KEY LABOR MARKET OUTCOMES AT 18 MONTHS AFTER RANDOM ASSIGNMENT, BY RURAL AND OLDER WORKER STATUS

	Rural	Older	Combined
Projected Sample Size	1,420	560	1,980
Likelihood of Starting a New Business	7.9	11.1	6.6
Likelihood of Employment	6.4	8.9	5.3
UI Benefit Amounts Received	$336	$471	$266
Total Earnings	$2,519	$3,529	$1,997

NOTE: The calculations are based on the same assumptions as those in Table 2. We also assume that the combined sample includes 1,420 urban and 560 rural workers.

TABLE 4

MINIMUM DETECTABLE EFFECTS FOR KEY LABOR MARKET OUTCOMES AT 18 MONTHS AFTER RANDOM ASSIGNMENT, BY SELECT PARTICIPANT CHARACTERISTICS

	Nonwhites	High School Diploma or Less	Prior Self-Employment Experience	Combined
Projected Sample Size	990	1,386	495	1,980
Likelihood of Starting a New Business	9.3	7.9	13.2	6.6
Likelihood of Employment	6.7	5.7	9.5	5.3
UI Benefit Amounts Received	$375	$317	$531	$266
Total Earnings	$2,518	$2,128	$3,564	$1,997

NOTE: The calculations are based on the same assumptions as those in Table 3. We also assume that the combined sample includes: 1) 990 nonwhites, 2) 1,386 workers with a high school diploma or less, and 3) 495 workers with prior self-employment experience.

The MDE results show that, based on the projected sample sizes, this study may detect statistically significant impact estimates for key participant subgroups. For example, as shown in Table 3, this study will detect whether the program's impact on the likelihood of starting a new business was at least 7.9 percentage points for rural dislocated workers and at least 11.1 percentage point for older dislocated workers. Moreover, this study will detect whether the program led to an increase in total earnings by $2,519 for rural and by $3,529 for older workers, respectively. Previous analyses of self-employment programs suggest that program impacts may exceed these minimum detectable effects. For example, Benus et al (1995) found impacts of the Self-Employment Assistance program on the percentage self-employed of 22 percentage points in Massachusetts and 12 percentage points in Washington, impacts on weeks employed of nine weeks in Massachusetts and five weeks in Washington, and impacts on annual earnings of $6,000 in Massachusetts and $300 in Washington. Except for the $300 earnings impact in Washington, we could detect these impacts even with smaller sample sizes. Similarly, the MDEs in Table 4 indicate that this study is likely to detect significant impacts on key outcomes for various subgroups, including nonwhites, individuals with no more than a high school diploma, and those with prior self-employment experience.

b. Unusual Problems Requiring Specialized Sampling Procedures.

There are no unusual problems requiring specialized sampling procedures.

3. Methods to Maximize Response Rates and Data Reliability

a. Response Rates

We expect to achieve an 80 percent response. In the follow-up surveys for Project GATE, which consisted of three waves of surveys at 6, 18, and 60 months after random assignment, IMPAQ achieved an 82% completion rate in Wave 1, 88% in Wave 2, and 81% in Wave 3[3]. Several strategies will be used to achieve this high response rate. First, before the survey begins, an advance letter describing the purpose and sponsorship of the survey will be mailed to potential respondents (the letter is presented in Appendix C). This advance letter will assure potential respondents that the caller is conducting a research survey and not soliciting donations or selling anything. Letters will be sent approximately one week before the sample is released to the CATI call scheduler. The letter will request up-to-date contact information and provide a toll-free call-in number.

Second, experienced interviewers will be recruited and extensively trained. These interviewers will be thoroughly trained on data collection procedures, including methods for promoting cooperation among sample members. Interviewers especially skilled at encouraging cooperation will be available to persuade reluctant respondents to participate and will be assigned to attempt conversions with respondents who initially refuse (except for hostile refusals). The survey will be translated into Spanish and bilingual interviewers will be used to conduct surveys in Spanish.

Third, call scheduling in CATI will allow respondents to select the time most convenient for them to be surveyed.

Fourth, detailed contact information provided by sample members will be used to help locate sample members for the follow-up surveys. This information will include not only an address and telephone number but also an e-mail address and cell phone number, if available, and the names and addresses of three relatives or other individuals who will know how to contact them. If these sources do not provide sufficient information to contact the sample member,

---

[3] http://wdr.doleta.gov/research/keyword.cfm?fuseaction=dsp_resultDetails&pub_id=2444&mp=y

extensive use will be made of various on-line databases to try to locate sample members who have moved.

Finally, a $15 incentive will be used to encourage participation. Past research (Singer 2002, Singer et al. 1999a and 1999b) indicate that incentive payments help boost survey response rates. They are particularly valuable in panel studies like this one where respondents are contacted and surveyed.

When the survey is completed we will conduct an analysis of non-response to assess whether the survey sample is representative of the initial population of GATE II applicants. In particular we will examine whether any differences in response rates between program and control group members may affect the findings. Quarterly wage record data on post random assignment earnings, not subject to non-response, will be used to examine differences in earnings. Sample weights will be assigned to adjust for differences between responders and non-respondents in important background characteristics.

Two kinds of bias arise from missing data in the surveys. "Non-contact bias" results from the failure to locate the respondent despite repeated telephone calls and other locating efforts. This may be for various reasons such as the respondent having died, moved, or become incarcerated. "Non-response bias" results from the respondent's failure to answer a particular question, either because the respondent refused to answer, or the respondent did not know the answer.

Adjusting for non-contact bias involves estimating a logistic regression model of the probability that a sample member responded to the survey and using the predicted probability of survey response to construct appropriate weights for each respondent. The initial logistic regression model will be estimated using all available applicants socioeconomic and prior employment characteristics. The model can be expressed by the following equation:

$P¿$

The dependent variable in this model (S) is the likelihood that the participant responds to the survey. S equals 1 if the participant responds to the survey and 0 otherwise. The initial model includes all available participant characteristics (X) collected by states through the program application form, including: gender, race, education, tenure with prior employer, prior wages, industry, and occupation. Inclusion of all available baseline characteristics in the model is

important for identifying which variables are strong predictors of non-response. To identify the final set of covariates for the non-response model (i.e., the variables that are strong predictors of non-response), we will use the following measures of predictive ability and goodness of fit:

- **McFadden's Pseudo R-squared statistic** – This measure captures the percent of the variation in the likelihood of responding to the survey that is explained by participant characteristics. This statistic is constructed as follows:[4]

$$R^2 = 1 - \frac{\ln\left(\hat{L}\left(M_{Full}\right)\right)}{\ln\left(\hat{L}\left(M_{Intercept}\right)\right)}$$

- **Akaike's Information Criterion (Akaike 1974)** – This statistic measures the efficiency of the model in predicting the outcome based on the number of covariates. This statistic is constructed as follows:[5]

$$AIC = 2k - 2\ln\left(\hat{L}_k\right)$$

These statistics will be used to assess which set of observable characteristics are strong predictors of survey non-response. Once we identify this set of characteristics, we will estimate the final version of the model using these characteristics. Once this model is estimated, we will use the results to calculate the predicted likelihood of survey response based on each participant's characteristics, as follows:

$$w_i = \frac{1}{f\left(X \cdot \hat{b}\right)}$$

In words, the survey non-response weight for each participant is the inverse of the predicted probability of response, $f\left(X \cdot \hat{b}\right)$. These weights will be used in the analyses to make the sample representative of all program applicants. This method is a widely accepted practice in program evaluations for controlling for survey non-response and for making estimation results representative of all program applicants (McConnell et al., 2006; Trenholm et al., 2007; Benus et al., 2009).

b.   Reliability of Data Collection

---

[4] Note: is the full model prediction using all control variables; is the full model prediction without predictors; and is the estimated likelihood of response based on the logit model.

[5] Note: is the number of parameters in the estimated model; and is the maximum value of the likelihood function based on the estimated model.

The draft questionnaire borrowed heavily from questionnaires developed for other U.S. Department of Labor studies, including *the Growing America Through Entrepreneurship (GATE) Demonstration* (OMB Number: 1205—0444)*; the *UI Self-Employment Demonstration Follow-up Survey*; *Comprehensive Assessment of Self-Employment Assistance Programs* (OMB Number 1205-0412); *1992 Economic Census Characteristics of Business Owners Survey* (OMB Number 0640-0022); the *Job Search Assistance Experiment Survey* (OMB Number 1205-0367), and the *Survey of UI Recipients* (OMB Number 1205-0405). The survey also drew from the non-federally funded study, the *Work First New Jersey Evaluation*. As a result, most of the questions in this survey have been thoroughly tested on large samples, all of which had prior OMB review and approval.

The questions were designed to ensure that they would be easily understood by respondents. Revisions were made to the draft questionnaire based on an internal review, a review by technical advisors to GATE II, a review by DOL, and a pretest.

The use of CATI to conduct the survey also helps ensure the reliability of the data. A CATI system controls question branching (reducing item nonresponse due to interviewer error), modifies wording (providing memory aids and probes and personalizing questions), and constructs complex sequences that are not possible to produce or are less accurate in hard-copy surveys. The probes, verifications, and consistency checks are built into the system; thus, the CATI system standardizes the procedures. These procedures ensure the reliability of the data collection methods and the data collected through those methods.

Lastly, IMPAQ International will monitor each interviewer's work using silent call-monitoring equipment and video monitors that display the interviewer's screen for real-time monitoring by supervisors.

4. Tests of Procedures or Methods

Pretest Procedures and Sample. The survey instrument in this OMB package is virtually identical to the 18-month (Wave 2) follow-up survey instrument that OMB approved for use by IMPAQ in Project GATE in 2005. IMPAQ conducted a pretest of that follow-up survey instrument using printed versions of the questionnaire. Each pretest was monitored to determine if the respondent's answers were consistent with the intent of the question. The pretest also enabled IMPAQ to check the accuracy of the internal skip patterns of the survey, the instructions

to interviewers, the ease with which interviewers could read question wording, and the apparent ability of respondents to follow the wording and sequence of questions. Project staff monitored all pretest surveys and took extensive notes on question wording, skip logic, and the overall flow of the instrument.

The pretest sample was drawn from two states that subsequently administered the GATE II demonstration – North Carolina and Virginia. We chose this sample to ensure some variety in program context and in personal experiences of respondents.

Pretest Results. As a result of the pretest, minor question-wording modifications and sequence changes were made to the instrument.

- In Section B, which deals with self-employment services, we removed a duplicate question and minor changes were made to clarify the skip logic for a few additional questions.

- In Section C, which deals with self-employment experience, minor changes were made to clarify the skip logic and how respondents were asked about their income.

- Minor wording changes were made in Section D, employment, working for someone else.

- In Section E, Income Sources and Amounts, questions were added to get a better understanding of respondents' unemployment compensation benefits.

- Minor wording changes were also made to Sections F and G.

Administration time in the pretest, adjusted for the expected efficiencies to be achieved in CATI operations, was found to be consistent with survey plans. Administration time in the pretest completions averaged 30 minutes. IMPAQ's actual experience administering the 18-month survey in Project GATE to over 3,000 respondents in 2005 through 2007 showed that the average time required to complete a survey was 25 minutes. The response time varied, depending on the specific skip patterns appropriate to the respondent's individual circumstances. The current survey instrument has several additional questions which increases the average time required from 25 minutes to 30 minutes per survey.

5. Individuals Consulted on Statistical Methods

The following persons outside of ETA contributed to, reviewed, and/or approved the design, instrumentation and sampling plan:

Name	Affiliation	Telephone Number
Jacob Benus (Project Director)	IMPAQ International	(443) 367-0088
Ted Shen	IMPAQ International	(443) 367-0088
Ben Hansen	IMPAQ International	(443) 367-0088
Terry Johnson	Battelle Memorial Institute	(206) 525 3130

REFERENCES

Agrest, Alan (1996). *An Introduction to Categorical Analysis*. Hoboken, New Jersey, John Wiley and Sons, Inc.

Akaike, Hirotsugu (1974). "A New Look at the Statistical Model Identification." *IEEE Transactions on Automatic Control*, 19(6).

Benus, Jacob M., Terry R. Johnson, Michelle Wood, Neelima Grover, and Theodore Shen. "Self-Employment Programs: A New Reemployment Strategy, Final Report on the UI Self-Employment Demonstration." Unemployment Insurance Occasional Paper No. 95-4. Washington, DC: U.S. Department of Labor, Employment and Training Administration, 1995.

Bloom, Howard S., (1995). Minimum Detectable Effects. Evaluation Review 19(5) 547-556.

Hosmer, D.W. Jr. and S. Lemeshow (1989). *Applied Logistic Regression*. New York, New York: John Wiley and Sons, Inc.

Little, R.J.A. and Rubin, D.B. (2002). *Statistical Analysis with Missing Data*. New York: Wiley.

Needels, Karen, Walter Corson, and Walter Nicholson. "Left Out of the Boom Economy: UI Recipients in the Late 1990s." ETA Occasional Papers 2002-2003, Washington, DC: U.S. Department of Labor, Employment and Training Administration, 2002.

Singer, Eleanor. "The Use of Incentives to Reduce Nonresponse in Household Surveys." In *Survey Nonresponse*. Edited by R. Groves, D. Dillman, J. Eltinge, and R. Little. New York: John Wiley and Sons, 2002.

Singer, Eleanor, John Van Hoewyk, Nancy Gebler, Trivellore Raghunathan, and Katherine McGonagle. "The Effect of Incentives on Response Rates in Interviewer-Mediated Surveys." *Journal of Official Statistics*, vol. 15, no. 2, 1999a.

Singer, Eleanor, Robert M. Groves, and Amy Corning. "Differential Incentives: Beliefs About Practices, Perceptions of Equity, and Effects on Survey Participation." *Public Opinion Quarterly*. Vol. 63, 1999b.

Walker, Britton, and Amy Kays Blair. *2002 Directory of U.S. Microenteprise Programs*. Washington, DC: The Aspen Institute, 2002.