

**Evaluation of the Pell Grant
Experiments Under the
Experimental Sites Initiative**

OMB Supporting Statement:

Part B

July 3, 2012

**Evaluation of the Pell Grant
Experiments Under the
Experimental Sites Initiative**

OMB Supporting Statement:

Part B

July 3, 2012

CONTENTS

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

 B1. Respondent Universe and Samples.....5

 B2. Statistical Methods for Sample Selection and Degree of Accuracy Needed..... 9

 B3. Maximize Response Rates..... 14

 B4. Tests of Procedures or Methods..... 17

 B5. Individuals Consulted on Statistical Aspects of the Design.....17

REFERENCES..... 18

TABLES

B.1 Survey Sample Sizes, by Experiment Using a Stratified Sampling Approach..... 8

B.2 Sample Sizes and Precision, by Experiment.....10

B.3 Sample Size and Minimum Detectable Impacts, by Experiment.....11

B.4 Survey Sample Sizes and Precision, by Experiment.....13

B.5 Survey Sample Size and Minimum Detectable Impacts, by Experiment..... 14

B.6 Schedule for Gaining Cooperation, by Type of Contact.....15

FIGURES

Figure B.1. Stylized Model of the Recruitment, Enrollment, and Random
Assignment Process for PGE When There Is Need-Blind
Admissions..... 3

Figure B.2. Time Line for the Pell Grant Experiments Study 7

PART B: COLLECTION OF INFORMATION EMPLOYING STATISTICAL METHODS

The Institute of Education Sciences (IES) at the U.S. Department of Education (ED) requests approval to conduct an evaluation of the effects of two Pell Grant Experiments (PGE) demonstrations under the Experimental Sites Initiative (ESI). The ESI, authorized by section 487A(b) of the Higher Education Act of 1965 (HEA), allows the Secretary to grant waivers from specific Title IV HEA statutory or regulatory requirements to enable institutions to test alternative methods for administering those federal student aid programs. The two demonstrations are targeted to income-eligible postsecondary students interested in vocational training but who could not otherwise receive a Pell grant because: (1) they currently have a bachelor's degree, or (2) they seek to enroll in a vocational program that is shorter than the current minimum duration and clock hours. Because of the potential high costs – and benefits – of expanding Pell grant eligibility in these two ways, ED has decided to rigorously assess the demonstration programs using a random assignment design. The study will examine the impacts of each experiment on employment and earnings, participation in education and training and job support activities, and student debt and financial aid receipt.

OVERVIEW OF THE DEMONSTRATIONS AND STUDY APPROACH

Under the ESI, Title IV institutions choose to participate in demonstrations or “experiments” in response to a notice from ED’s Office of Federal Student Aid (FSA). FSA published such a notice in October 2011, inviting postsecondary schools to participate in any of 8 different experiments¹, two of which expanded Pell grant eligibility for students seeking job training. That notice also specified the institutions’ obligations to provide data and to ensure that a control or comparison group could be formed so that the effects of participating in the experiments could be evaluated. In subsequent webinars, FSA has provided additional detail to interested institutions about the demonstrations and the evaluation.

1. The Two Pell Grant Experiments (PGE)

Under the current ESI, postsecondary schools will receive waivers to enable them to provide Pell Grants to students who would not otherwise qualify under current Pell Grant rules. The PGE evaluation will include two substudies, each of which relaxes one eligibility criterion for receipt of a Pell Grant:

¹ See <https://experimentalsites.ed.gov/exp/index.html>

1. **Experiment 1.** Students who already hold a bachelor's degree and who document that they are unemployed or underemployed will be able to receive Pell Grant award support. This support can be for up to a one-year program of vocational education intended to help them obtain employment, to be used over no more than two award years. Current rules do not allow individuals with a bachelor's degree to receive Pell support unless it is to be used for teacher certification or licensure.
2. **Experiment 2.** Students will be able to receive a prorated amount of Pell Grant financial support for short-term vocational training that lasts for at least 150 clock hours over a period of at least 8 weeks. Current rules require that a student's academic program is at least 600 clock hours (or an equivalent in semester, trimester, or quarter hours) over at least 15 weeks to qualify for Pell support.

2. Selecting Schools

Schools that volunteered to implement Experiments 1 and 2, that were in good standing in administering Title IV programs (e.g., related to compliance, default rates, etc.), and that agreed to meet the requirements of the evaluation form the study school sample. ED expects the sample to include a maximum of 28 schools for Experiment 1 and 40 for Experiment 2, but with approximately 17 intending to participate in both experiments. Although there will be 51 distinct schools participating, because each experiment will be studied separately there will be a total of 68 experiments underway. Each school will identify the set of vocational or job training programs to which the experiments will apply.

3. Identifying Eligible Students

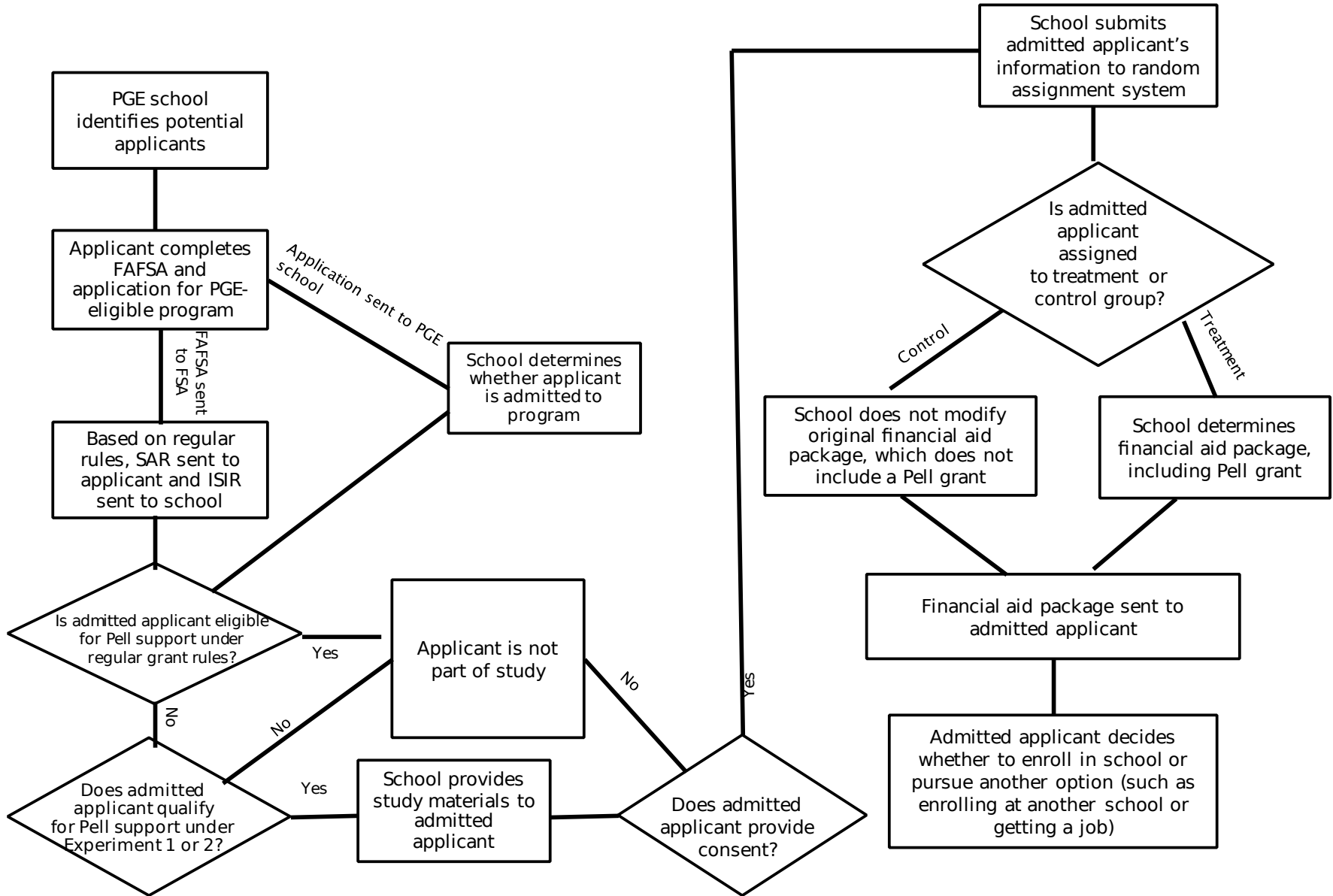
Recruitment, enrollment, and random assignment of sample members into the PGE study will be the same for both substudies and will involve several steps (Figure A.1). Participating schools will recruit applicants and encourage them to submit both the Free Application for Federal Student Aid (FAFSA) (typically completed on line) and an application to the PGE-eligible program in which the student wants to enroll. Simultaneously or sequentially, FSA will process the FAFSA and the school will determine whether the student can be admitted to the vocational program. Students will receive a Student Aid Report (SAR) and schools an Institutional Student Information Record (ISIR), which provides an assessment of the applicant's expected family contribution (EFC) towards his or her educational expenses.

Because the potential participants in the study would not ordinarily be eligible for Pell grants, by virtue of their educational characteristics or their program, the PGE schools will need to determine a way to identify candidates for the experiments rather than processing their aid packages in

OMB Package for the Pell Grant Experiments

the usual manner. Most likely, the institutions will ensure that financial aid office staff flag students who apply to the PGE eligible programs and review their ISIRs separately.

Figure B.1. Stylized Model of the Recruitment, Enrollment, and Random Assignment Process for PGE When There Is Need-Blind Admissions



4. Random Assignment

Once candidates for the experiments are identified by the institutions, these eligible individuals will receive information about the study and be asked to provide consent for their participation. School staff will data-enter into a web-accessible, study-specific random assignment system the names and Social Security numbers of eligible admitted applicants who have given consent, as well as a very limited amount of other information about the individual and PGE program, so that random assignment can be conducted.² In real-time (with little delay), the school then will be notified of the research group status of each study participant. The proportion assigned to the treatment group versus control group will depend on the number of eligible candidates the institutions expect to identify.

Control group members will have access to the normal financial support that they are eligible for (i.e., excluding a Pell Grant). Study participants assigned to the treatment group will be offered a Pell grant, and the school will take this into account in determining any other aid for which the student is eligible. The financial aid packages will then be provided to the study participants. Regardless of whether the participant is assigned to the treatment or control group, he or she can choose to enroll at the PGE school, enroll at another school to which he or she has been admitted, or pursue some other type of activity.³

It is expected that schools in Experiment 1 will enroll 100 participants, on average, while schools in Experiment 2 will enroll 200 participants into the study, for a total of 2,800 sample members in Experiment 1 and 8,000 in Experiment 2. Thus, total sample enrollment for the study will be 10,800. The study participants will consist of individuals who have been determined to be eligible for the study under either Experiment 1 or Experiment 2 and who have consented to be in the study.

² If possible, we would like to randomly assign within program area to ensure treatment-control group balance on this important dimension. This might allow the evaluation to calculate impacts separately by occupational area.

³ The particular methods that schools use to recruit potential sample members and any screening that is conducted to assess applicants' interest levels in the PGE program before random assignment is conducted will have an influence on the rate at which study participants enroll in the PGE program.

5. Collecting Data

Both substudies of PGE will have the same data collection plans. These collection plans include new burden imposed by two types of data collection efforts: (1) PGE school data for all study participants and (2) survey data for a subsample of 2,000 participants (out of 2,500 participants randomly selected for participation in the survey). The plans also include use of two other types of data—FSA data and annual earnings data maintained by the Social Security Administration (SSA)—that do not generate data collection burden on participating schools or students. These data are described in detail in Section A.2. Together, these data will provide a rich set of information from which we can estimate the impacts of expanded Pell grant eligibility on study participants' employment and earnings, educational experiences, and student debt, as well as on the characteristics of participants and their vocational programs.

6. Reporting

The schedules for sample enrollment and program participation, as well as when post-program outcomes can be observed, drives the project's reporting schedule. The study is expected to last five years, from spring 2012 to June 2017 (Figure A.2)⁴. Enrollment of school applicants into the study is expected to start in summer 2012. Although each of the 68 experiments in the study might take a slightly different amount of time to complete its enrollment of study participants, enrollment for the study is expected to continue through spring 2014.

Most of the study participants who enroll in Experiment 2 are expected to complete their participation in education or training in a fairly short time (two to four months), while participants who enroll in Experiment 1 are expected to take 9-14 months but could be up to two years if attending less than full-time. It is expected that all sample members who participate in a PGE program will complete their training program by late 2014. The first full post-program calendar year for all study participants will be 2015, although many of the participants who entered the study early in the sample enrollment period are expected to have had a full year of post-program experiences prior to then. SSA data covering calendar year 2015 is expected to be available for analysis in preliminary form in spring/summer 2016,

⁴ While Ed would prefer to have an additional, longer-term follow up on earnings, study resources do not currently allow for that.

making it possible to draft a report and have it go through IES' statutorily required review process for publication in late spring 2017.

B1. Respondent Universe and Samples

Of the four data collection efforts, three will provide administrative data for all study participants: the FSA data, the PGE school data, and the SSA data. As noted earlier, the FSA data and the SSA data will not generate new burden as a result of the study. Nevertheless, the discussion in this section groups these three sources of data together, because of their similarities. Furthermore, a subsample of 2,500 participants, out of an expected 10,800, will be asked to take part in the survey. The study aims to have 2,000 participants respond to the survey. The next two subsections describe the respondent universe and samples for the administrative data and the survey of study participants.

a. Administrative Data

The study is designed to collect data on individuals who are ineligible for the Pell Grant program because they either (Experiment 1) applied to vocational or career training programs but already have a bachelor's degree or (Experiment 2) applied to a short-term training program. In spring 2012, the Office of Federal Student Aid (FSA) recruited schools to volunteer programs for the study. As described earlier, FSA expects 28 schools to participate in Experiment 1 and 40 schools to participate in Experiment 2. On average, each school in Experiment 1 will recruit 100 participants, and each school in Experiment 2 will recruit 200 participants. The potential respondent universe of respondents consists of these 10,800 study participants with 2,800 in Experiment 1 and 8,000 in Experiment 2. The data collection effort is designed to be representative of the two groups of individuals at the programs in the PGE study. It does not generalize to any other population of individuals or programs due to the process used to select schools (open invitation plus screening for Title IV administrative compliance by FSA), programs (the criteria listed in the invitation notice plus schools' preferences), and students (recruiting approaches used by schools).

All three types of administrative data are expected to be comprehensive in their coverage. Data on eligible candidates entered by PGE school staff for the purpose of random assignment will define the universe of study participants. The evaluation contractor will request data extracts from PGE school records for this sample of potential and actual enrollees; study

participants without an enrollment record are assumed to not have enrolled in a program at a PGE school. It is assumed that PGE programs already track student enrollment because they must verify it before students can receive financial aid.⁵ In addition, PGE programs are already required to report six-year graduation rates to the Integrated Postsecondary Data Education System. As a result, it is likely that PGE programs already have databases that track graduation outcomes over time.

The study expects a 100 percent response rate for the PGE program data collection effort. The federal notice inviting schools to participate in the experiments and subsequent communication from FSA requires that all PGE schools provide relevant administrative data as a condition for participation. As a result, the study will include only individuals with administrative data from PGE programs.

The FSA and SSA data also are expected to be available for 100 percent of study participants. The cause of the 100 percent response rate of the FSA data is analogous to that of the PGE program's data. In both cases, only individuals with the data are eligible to participate in the study. The study will assume that individuals without an SSA earnings record have zero earnings and no employment. This approach is consistent with that of others studies that use data from the SSA Master Earnings File (Schochet et al. 2003).

⁵ The characteristics of study participants can be collected by the school that houses the PGE program.

OMB Package for the Pell Grant Experiments

Figure B.2. Time Line for the Pell Grants Experiments Study

Contract Month	2012			2013				2014				2015			
	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec	Jan-Mar	Apr-Jun	Jul-Sep	Oct-Dec
1. PGE School Planning Phase	▬														
2. Study Participant Enrollment		▬	▬	▬	▬	▬	▬	▬							
3. Expected Program Completion Period			▬	▬	▬	▬	▬	▬	▬	▬	▬				
4. FSA Data Extraction			◇				◇				◇				
5. Data Extracts from PGE Schools				◇				◇				◇			
6. Analysis of PGE School Data												▬	▬		
7. Analysis of SSA Data															
8. Survey Fielding Period										▬	▬	▬	▬		
9. Analysis of Survey Data												▬	▬	▬	▬
10. Final Report															

△ Draft deliverable. ▲ Final deliverable. ◇ Data extract.

b. Data from the Survey of Study Participants

The second data collection effort will be a survey of participants. The size of the sample was chosen to provide useful insights while fitting within the available resources of the study. As shown in Table B.1, 2,500 study participants will take part in the survey. This size implies an overall sampling rate of 23.1 percent. It uses a stratified sampling method with four strata: (1) the treatment group in Experiment 1; (2) the control group in Experiment 1; (3) the treatment group in Experiment 2; and (4) the control group in Experiment 2.

Table B.1. Survey Sample Sizes, by Experiment Using a Stratified Sampling Approach

Stratum	Number of Individuals in the Sampling Frame	Sampling Rate	Number of Individuals that Take Part in the Survey	Survey Response Rate	Number of Survey Respondents
Experiment 1					
Treatment group	1,867	33.5%	625	80%	500
Control group	933	67.0%	625	80%	500
Experiment 2					
Treatment group	5,333	11.7%	625	80%	500
Control group	2,667	23.4%	625	80%	500
Total	10,800	23.1%	2,500	80%	2,000

Not all of the 2,500 individuals who take part in the survey will respond to it. Based on prior studies of a similar population (McConnell et al. 2006), the study will aim to achieve a response rate of 80 percent across all strata.

Given this target response rate, it is expected that there will be 2,000 individuals for whom survey data will be available. Experiments 1 and 2 will each have 1,000 survey respondents. As discussed in further detail in Section B.2, a nonresponse analysis will be conducted and, as needed, the survey weights will be adjusted to take into account the probabilities of response by different types of sample members.

ED is exploring different approaches to capitalize on cost efficiencies that would occur with having a relatively short survey fielding period.⁶ The different approaches have implications for the generalizability of the survey results. One approach under consideration is to select survey participants from an early subset of the time window during which sample members enrolled in the study. This approach could allow for a uniform follow-up period across survey participants but the survey results would then generalize to the study participants from the enrollment period in which the survey sample was drawn and not to all study participants. Another approach would be to randomly select for the survey study participants from the full enrollment period; if this approach is taken, then the lengths of follow-up would vary across survey respondents. However, the survey sample would be representative of all study participants. The pros and cons of these approaches will be considered once an evaluation contractor has been selected.

B2. Statistical Methods for Sample Selection and Degree of Accuracy Needed

The study will apply the proper statistical methods to generate rigorous answers to the research questions. These methods pertain to the sampling frameworks and estimation procedures. The choice of method is based on the nature of the analytical sample, which will be either (1) all study participants or (2) the sample of study participants that take part in the survey. The next subsections describe the statistical methods used for these two analytical samples.

a. Sampling Methods and Analysis of Data for All Study Participants

The administrative data from FSA, PGE schools, and SSA will contain information for all study participants.⁷ The analytical sample is designed to generalize to the universe of individuals eligible for the Pell grant experiments demonstrations who would otherwise be ineligible for Pell

⁶ The approach taken is not expected to affect the estimates of the burden created by the survey effort and reported in this package.

⁷ If a match for a sample member is not found, it will be assumed that he or she did not participate in the activity covered by the data.

Grants at the PGE programs. Because the analytical sample will have data on 100 percent of the study participants, the study will not need to use sampling weights to correctly represent the population. Implicitly, the sampling weight for each respondent will be one.

To demonstrate the precision associated with this sampling approach, Table B.2 provides the half-widths of the 95 percent confidence intervals for two potential proportions of the outcome variable. For Experiment 1, the half-width of the confidence interval is 0.027 when half of the population has a particular outcome. When the proportion is only 0.10, the half-width falls to 0.016. In addition, the half-width of the confidence interval for annual earnings is \$758. The confidence intervals for Experiment 2 are smaller than those of Experiment 1 because it has a larger sample size. These figures indicate that the study will produce relatively precise estimates of the outcome variables for both experiments.

The study will also produce descriptive statistics by treatment and control group within each experiment. Even with these smaller sample sizes, the study will lead to relatively precise estimates of the outcome variables. For example, the half-width of the confidence interval for a proportion of 0.50 for the control group in Experiment 1 is 0.046. The corresponding confidence interval for earnings is \$1,313. These two half-widths represent the least precision available to the study using the PGE program data. Even so, the estimates from this sample are will provide useful insights about the population of study participants.

Table B.2. Sample Sizes and Precision, by Experiment

	Sample Size	Half-Width of Confidence Interval of a Proportion of 0.50	Half-Width of Confidence Interval of a Proportion of 0.10	Confidence Interval of Earnings (dollars)
Experiment 1	2,800	0.027	0.016	758
Treatment group	1,867	0.033	0.020	929
Control group	933	0.046	0.028	1,313
Experiment 2	8,000	0.018	0.011	498
Treatment group	5,333	0.021	0.013	610
Control group	2,667	0.030	0.018	862

Note: The confidence intervals are based on a 95 percent probability level. The intraclass correlation is equal to 0.04. The confidence intervals are based on the effective sample size, which is equal to the sample size divided by the design effect.

The estimation procedures used for this analytical sample are designed to measure the impacts of the offer of Pell Grants. Because the average Pell Grant amount, program content/duration, and student characteristics will differ by experiment, the study will analyze the impacts separately for each experiment. The study will estimate ordinary least squares regression models in the form of Equation (1). The dependent variable is y_{ip} , where y is the outcome of interest for study participant i in program p . The main outcome variables will be employment and earnings, but the study will also include enrollment, graduation, and other measures as secondary outcome variables. The variable g_i indicates whether the study participant was randomly assigned to be in the treatment or control group. This specification implies that the parameter γ is the effect of access to a Pell Grant on the outcome y . In this setup, γ is the average treatment effect of the Pell Grant access for this population.

$$(1) \quad y_{ip} = X_{ip}\beta + \gamma g_i + \mu_p + \varepsilon_{ip}$$

The regression model will control for a variety of characteristics in X_{ip} , such as the participant's age, educational background, and earnings before random assignment. The inclusion of the control variables will enable the study to estimate the effects of Pell Grants with a high degree of precision. The remaining terms μ_p and ε_{ip} represent program fixed effects and a stochastic error term, respectively.

To determine whether the study can detect the impact of Pell Grants, Table B.3 presents the minimum detectable impacts (MDIs) of the estimation procedure, which are defined as the minimum detectable effects times the standard deviations of the outcomes. The power calculations are based on a

two-tailed test with an alpha of 0.05 and a beta of 0.80. They are also based on an assumed intraclass correlation of 0.04 at the school level. The mean and standard deviation of enrollment and completion are based on public statistics on two-year colleges (Snyder and Dillow 2011). The mean and standard deviation of earnings and employment are based on a prior study of adults seeking job training assistance (Bloom et al. 1993).

Table B.3. Sample Size and Minimum Detectable Impacts, by Experiment

	Sample Size	Mean	Standard Deviation	MDI with R ² =0.2	MDI with R ² =0.4
Experiment 1					
Employment	2,800	0.8	0.4	0.058	0.050
Earnings	2,800	\$10,436	\$14,198	\$2,056	\$1,780
Enrollment	2,800	0.7	0.5	0.066	0.057
Completion	2,800	0.5	0.5	0.072	0.063
Experiment 2					
Employment	8,000	0.8	0.4	0.038	0.033
Earnings	8,000	\$10,436	\$14,198	\$1,349	\$1,169
Enrollment	8,000	0.7	0.5	0.044	0.038
Completion	8,000	0.5	0.5	0.048	0.041

Notes: The power calculations are based on an alpha of 0.05 and a beta of 0.80. The MDIs are for differences between the treatment and control groups, where the treatment group is two-thirds of the sample and the control group is one-third of the sample. The results are based on a 100 percent response rate to the administrative data. The intraclass correlation is set equal to 0.04. The power calculations are based on the effective sample size, which is equal to the sample size divided by the design effect.

MDI = minimum detectable impact.

Under standard assumptions, the power calculations show that the estimation procedure can detect meaningful differences between the treatment and control groups.⁸ For example, with an R² equal to 0.2, the procedure is powered to detect a difference of 5.8 percentage points in the probability of employment and a difference of \$2,056 in earnings in Experiment 1. Both experiments are powered to detect even smaller differences when the regression model explains a larger portion of the variance in the outcome (R² = 0.4). In this setting, the procedure is powered to detect a difference of 5.0 percentage points in the probability of enrollment and a difference of \$1,780 in earnings in Experiment 1. Thus, the

⁸ The table presents hypothetical means and standard deviations that could be expected based on other research. The particular methods that schools use to recruit potential sample members and any screening that is conducted to assess applicants' interest levels in the PGE program before random assignment is conducted will have an influence on the rates at which treatment and control group members participate in an educational program and achieve other outcomes of interest to the study. A lower rate of enrollment than is assumed in the table, for example, would lead to a higher MDI.

estimation procedures are likely to detect the true effects of access to Pell Grants on the outcomes if the true effects exceed these MDIs.

These MDIs are near the estimated impacts found in two recent studies of employment and training programs, although some caution is warranted in drawing conclusions. The Sectoral Employment Impact Study (SEIS), a random assignment study of an intervention for underskilled, unemployed, and low-income adults, found impacts on earnings in the second year of follow-up of about \$4,000 per year (Maguire et al. 2010). SEIS examined three programs that offered a combination of short-term training and job placement assistance for unemployed and low-income adults. Another study, the Workforce Investment Act Non-Experimental Net Impact Evaluation (Heinrich, Mueser, and Troske 2009), found a difference between a treatment group and a comparison group that equated to about \$1,800 for men and \$2,600 for women on an annual basis. This study examined the effects of Workforce Investment Act services on dislocated workers and adults who were generally low-income. Caution is warranted, however, because the latter study was non-experimental and there is a possibility that the estimated impacts are larger than what would have been found under an experimental study design. In addition, an evaluation of training programs for disadvantaged adults in the late 1980s, the National Job Training Partnership Act Study, found impacts that are roughly equivalent to \$608 in annual earnings for men and \$840 in annual earnings for women in 2010 dollars. Finally, it is unclear how differences between the interventions examined through these three studies and the PGE intervention, as well as differences in the populations served by the interventions, might influence the magnitude of expected impacts.

b. Sampling Methods and Analysis of Data from the Survey

The survey data will contain information for only a random subsample of study participants. As shown in Section B.1, the study will identify a stratified sample of study participants. It will use a disproportionate allocation in which the two experiments have the same expected number of individuals that take part in the survey. Each experiment will also have an equal number of treatment and control individuals that take part in the survey. The choice of a disproportionate allocation is based on the goal of improving the precision of survey-based estimates for certain strata. The sampling rate for Experiment 1 (33 percent for the treatment group and 67 percent for the control group) is disproportionately large because it has fewer study participants than Experiment 2. The sampling rates for the two control groups (67 percent for Experiment 1 and 23 percent for Experiment 2) are also disproportionately large because there are fewer study participants in the control groups than the treatment groups. These sampling rates increase

the precision of the estimates for study participants who were not given access to a Pell Grant, because a key purpose of the survey is to understand their access to education and training without such aid.

Based on an 80 percent response rate, each experiment will have a sample of 1,000 survey respondents; the treatment and control groups will each have 500 respondents. The study will apply inverse probability weights to correctly represent the population of interest.

Given this expected response rate, the study will conduct a nonresponse analysis and make the appropriate nonresponse adjustments. In essentially all surveys, the sampling weights need to be adjusted to account for survey participants who cannot be located or who refuse to respond when located. The study will estimate logistic regression models that analyze the probability of responding to the survey. Based on these statistical models, the study will use the inverse of the propensity score as the adjustment factor. The study will then construct a new sampling weight that is the product of the original sampling weight and the adjustment factor. By making this nonresponse adjustment, the sample will produce unbiased estimates of population parameters under certain statistical assumptions.

To demonstrate the precision associated with this sampling approach, Table B.4 provides the half-widths of the 95 percent confidence intervals for two potential proportions of the outcome variable. The half-width of the confidence interval is 0.045 when half of the Experiment 1 sample has a particular outcome. When the proportion is only 0.10, the half-width falls to 0.027. In addition, the half-width of the confidence interval for annual earnings is \$1,269. The confidence intervals for Experiment 2 are slightly larger than those of Experiment 1 because they represent a larger universe of study participants (8,000 participants compared with 2,800 for Experiment 1). These figures indicate that the survey will produce estimates of the outcome variables that are less precise than those from administrative data on all study participants.

Table B.4. Survey Sample Sizes and Precision, by Experiment

	Sample Size	Half-Width of Confidence Interval of a Proportion of 0.50	Half-Width of Confidence Interval of a Proportion of 0.10	Confidence Interval of Earnings (dollars)
Experiment 1	1,000	0.045	0.027	1,269
Treatment group	500	0.063	0.038	1,795
Control group	500	0.063	0.038	1,795
Experiment 2	1,000	0.050	0.030	1,408
Treatment group only	500	0.070	0.042	1,991

OMB Package for the Pell Grant Experiments

Control group only	500	0.070	0.042	1,991
--------------------	-----	-------	-------	-------

Note: The confidence intervals are based on a 95 percent probability level. The intraclass correlation is set equal to 0.04. The confidence intervals are based on the effective sample size, which is equal to the sample size divided by the design effect.

Because the study will use survey data to learn about the financial aid and job training activities for the treatment and control groups, Table B.4 also presents the precision estimates by experimental group. With these smaller sample sizes, the study will obtain less precise estimates of the outcome variables. For example, the half-width of the confidence interval for a proportion of 0.50 for the treatment group in Experiment 2 is 0.070. The corresponding confidence interval for earnings is \$1,991. The half-widths for the control group are identical to those of the treatment group; the estimates in Experiment 1 will be more precise than those of Experiment 2.

The estimation procedure for analyzing the survey data is analogous to those for all study participants. To determine whether this approach is able to detect the effect of Pell Grant accessibility, Table B.5 presents the MDIs for each outcome. The fifth column shows the MDIs when the regression model has an R-squared of 0.2. The rightmost column shows the MDIs when the R-squared is equal to 0.4. When the model explains a greater fraction of the variation in the outcome variable, the specification is able to detect smaller effects of access to a Pell Grant.

In general, the MDIs using the survey data are larger than those using the administrative data and what might be expected base on results from other research studies about employment and training interventions. Given this situation, the study will view the survey results as being exploratory in nature.

Table B.5. Survey Sample Size and Minimum Detectable Impacts, by Experiment

	Sample Size	Mean	Standard Deviation	MDI with R ² =0.2	MDI with R ² =0.4
Experiment 1					
Employment	1,000	0.8	0.4	0.091	0.079
Earnings	1,000	\$10,436	\$14,198	\$3,243	\$2,809
Enrollment	1,000	0.7	0.5	0.105	0.091
Completion	1,000	0.5	0.5	0.114	0.099
Experiment 2					
Employment	1,000	0.8	0.4	0.101	0.088
Earnings	1,000	\$10,436	\$14,198	\$3,598	\$3,116
Enrollment	1,000	0.7	0.5	0.116	0.101
Completion	1,000	0.5	0.5	0.127	0.110

Notes: The power calculations are based on an alpha of 0.05 and a beta of 0.80. The MDIs are for differences between the treatment and control groups, where both the treatment and control groups are half of the sample. The results are based on an 80 percent response rate to the survey. The intraclass correlation is equal to 0.04. The power calculations are based on the effective sample size, which is equal to the sample size divided by the design effect.

MDI = minimum detectable impact.

B3. Maximize Response Rates

As explained in Section B.1, it is expected that the study team will be able to attain FSA, PGE school, and SSA data for all study participants. The collection and analysis of these data will be based on the assumption that there is a 100 percent match rate between the list of study participants and the administrative data records files. If a study participant is not in the SSA data files, for example, it will be assumed that he or she did not have Social-Security-covered earnings during the relevant time period. Therefore, the only data collection effort for which achieving a high response rate could be especially challenging is the follow-up survey for individuals. As a result, the discussion here focuses on the strategies the study team plans to use to ensure that it achieves a high survey response rate—strategies that have been successfully used in other studies.

Contact with sample members. The evaluation team will send an initial invitation letter on ED letterhead to sample members. This letter will (1) introduce the study and its purpose; (2) highlight ED as the study sponsor; (3) explain the voluntary and confidential nature of participation; (4) extend the incentive offer; (5) provide web survey log-in information; and (6) give a toll-free number for respondents to call in for questions or if they want to complete the survey by telephone. The envelope will be printed with the ED logo to capture the sample members’ attention and to communicate the legitimacy of the study. The contractor’s return address will be used to facilitate the processing of returned mail and locating procedures. The

advance mailing will include an information sheet providing answers to questions that sample members might have about the study. It also will include a telephone number and an ED website address that sample members can use to learn more about the study. Timed reminders offering the option to complete the survey via the telephone, paper, or web will follow the initial invitation letter.

Before the mailing of these materials, interviewing staff will be thoroughly trained on how to address respondents' questions about the study and questionnaire. In addition to the sheet of answers to questions that will accompany the advance mailing, a list of frequently asked questions and answers (FAQs) will be developed for the interviewers' use. The operational procedures manual for the computer-assisted telephone interviewing (CATI)-administered questionnaire will include these FAQs. The FAQs will also be available online for the self-administered web survey and web survey respondents will have access to them throughout the survey.

Locating sample members. A key component to obtaining a high response rate is locating sample members. The process of locating study participants will occur each time the study team collects administrative data from the PGE programs. This locating process will involve the use of an independent vendor that will check the full sample against current address databases. This first step is critical given that some sample members could have moved since they completed their FAFSAs, which is the initial source of locating information in the study data. The study team will use extensive tracking and locating procedures that have proven successful in other studies for sample members whose mail is returned as undeliverable. These include using other independent databases, checking with neighbors and family members, and searching social networking sites. When talking with contacts, the specific purpose of the study will not be disclosed, but it will be stated that the effort to reach the sample member is for an important study being sponsored by the government.

Gaining and maintaining cooperation. A second key component to achieving high response rates is gaining cooperation after locating respondents (Table B.6). Sample members who are difficult to contact and who have not yet completed the survey on the web will receive a reminder letter one week after the initial invitation letter and another reminder letter along with a paper copy of the questionnaire three weeks after initial contact. Reminder calls/interviews will begin four weeks after data collection starts for each sample member. Additional contacting efforts will continue through the end of the data collection period for remaining nonrespondents. To those sample members who refuse to participate, a targeted refusal conversion letter that will address their specific concerns will be mailed first.

Next, expert refusal conversion interviewers will make follow-up calls to try to gain the sample members' cooperation.

Table B.6. Schedule for Gaining Cooperation, by Type of Contact

Week	Type of Contact
0	Initial invitation letter (includes web log-in and password information)
1	Reminder letter (includes web log-in and password information)
3	Reminder letter (includes web log-in and password information, as well as paper copy of the questionnaire); refusal conversion letter is mailed
4	Reminder calls/interviews; refusal conversion begins
5, 7	Additional reminder material will be mailed and or calls will be conducted

Multiple language survey administration. During telephone contact, interviewers will identify Spanish-speaking respondents and connect them to or schedule them to speak with a bilingual interviewer. When necessary, translators for languages other than Spanish will be used.

Incentives for survey participants. Offering an incentive for the follow-up survey is essential to generate the desired response rates and reduce overall survey costs without affecting data quality. There is substantial evidence on the benefits of offering incentives. According to Singer et al. (2000), incentives can help achieve high response rates by increasing the sample members' propensity to respond; by doing so, 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. Incentives also increase the likelihood of participation from subgroups with a lower propensity to cooperate with the survey request. This is an important component of ensuring the representativeness of the survey respondents and the quality of the data collected. For example, Jäckle and Lynn (2007) found that incentives increased the participation of sample members more likely to be unemployed. There is also evidence that incentives bolster participation among those with lower interest in the survey topic (Schwartz et al. 2006; Jäckle and Lynn 2007; Kay 2001), resulting in data that are more nearly complete. Furthermore, paying incentives does not impair the quality of the data obtained (such as item nonresponse or the distribution of responses) from groups that would otherwise be underrepresented in the survey (Singer et al. 2000).

An incentive will be offered to all survey respondents, using a two-tiered incentive offer to encourage the selection of the less-expensive web option for survey administration—\$15 for completion on the web and \$10 for completion using CATI or on paper. It is anticipated that a substantial number of sample members will choose the web, because many of them are likely to be more comfortable with this self-paced, self-administered approach. Also, the higher incentive offer for web completion will encourage many to use that option. The web survey will be available as soon as invitations are mailed to sample members. It is estimated that 40 percent of the completed surveys will come from the web. However, survey participants will be offered the opportunity to complete the survey either through the telephone or by mail if they prefer—further boosting potential cooperation levels.

To leverage fully the benefits of offering incentives in the PGE evaluation, the advance letter to the study participants will mention the incentive. Interviewers will also mention the proposed incentive when they establish contact with the participants and attempt to gain their cooperation.

Survey length. The follow-up survey questionnaire is designed to be easy to complete. The questions are written in clear and straightforward language. The average time required for the respondent to complete the survey is estimated at 15 minutes.

Interviewer training. All contractor staff assigned to the study will participate in general training as well as project-specific training. The project-specific training will include role playing with scenarios and other techniques to ensure that interviewers can respond effectively to sample members' questions. The training will review responses to FAQs and each questionnaire item. Training sessions will stress the importance of being sensitive to respondents' situations while remaining impartial. The sessions will also focus on developing skills for securing respondents' cooperation and averting and converting refusals.

Targeted response rate and weighting. Employing these procedures, an 80 percent response to the survey is targeted. When the survey is completed, an analysis that compares respondents with nonrespondents will be conducted to assess whether the survey sample is representative of the target population of PGE participants. This analysis will use key variables available for all sample members through the FAFSA and other administrative data. If it appears that the survey respondent sample is not representative, sample weights will be adjusted for nonresponse.⁹ Response

⁹ Because the survey sample was selected using a stratified approach with different sample selection probabilities for each stratum, sampling weights will be applied to the

weights will be generated for subgroups, the characteristics of respondents and nonrespondents will be compared, and factors that explain nonresponse will be used to generate nonresponse weight adjustments.

B4. Tests of Procedures or Methods

The process to develop the survey instrument (provided in Appendix A) has drawn from previously used items, including many from prior research studies that have focused on helping individuals participate in education and training and achieve good employment and earnings outcomes. Therefore, the pretests of the instrument are expected to focus on ensuring that the question flow works well and that the time required for a respondent to complete the instrument is accurately estimated. The instrument will be pretested with a convenience sample of nine or fewer individuals. The pretest will be conducted iteratively, in two stages, so obvious improvements to the instrument will be incorporated before subsequent pretests are conducted. To avoid interviewer effects, the pretests will be conducted using more than one interviewer.

The pretests will be conducted by the evaluator once a contract is awarded (expected August 2012). Should the pretests result in any recommendations for changes to the estimated respondent burden or the instrument, we will notify OMB and request a formal revision through a change sheet.

B5. Individuals Consulted on Statistical Aspects of the Design

The study is based on the best possible decisions for the statistical aspects of the design. In doing so, it will provide rigorous answers to the research questions that will be of use to ED. During the study, it is expected that ED will consult with the evaluation contractor that is selected to assist ED in the completion of the study.

survey data to account for this differential probability of selection even if it is determined that there is no nonresponse bias.

REFERENCES

- Bloom, Howard et al. "The National JTPA Study: Title II-A Impacts on Earnings and Employment at 18 Months." Bethesda, MD: Abt Associates, Inc., 1993.
- Heinrich, Carolyn, Peter R. Mueser, and Kenneth R. Troske. "Workforce Investment Act Non-Experimental Net Impact Evaluation: Final Report." Washington, D.C.: U.S. Department of Labor, Employment and Training Administration Occasional Paper 2009-10, 2009.
- 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.
- Kay, Ward R. "The Use of Targeted Incentives to Reluctant Respondents on Response Rates and Data Quality." *Proceedings of the American Association for Public Research*. Montreal, Canada: American Association for Public Opinion Research, 2001.
- Maguire, Sheila, Joshua Freely, Carol Clymer and Maureen Conway. "Turning in to Local Labor Markets: Findings from the Sectoral Employment Impact Study." Philadelphia, PA: Public/Private Ventures, July 2010.
- Schochet, Peter Z., Sheena McConnell, and John Burghardt. "National Job Corps Study: Findings Using Administrative Earnings Records Data." Final report prepared for the U.S. Department of Labor. Princeton, NJ: Mathematica Policy Research, October 2003.
- Schwartz, Lisa K., Lisbeth Goble, and Edward M. English. "Counterbalancing Topic Interest with Cell Quotas and Incentives: Examining Leverage-Saliency Theory in the Context of the Poverty in America Survey." *Proceedings of the American Association for Public Research*. Montreal, Canada: American Association for Public Opinion Research, 2006.
- 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.
- Snyder, Thomas D., and Sally A. Dillow, S.A. "Digest of Education Statistics 2010." NCEES 2011-015. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, April 2011.

