

NCER-NPSAS Grant Study

Financial Aid Nudges 2017: A National Experiment to Increase Retention of Financial Aid and College Persistence

Supporting Statement Part B

OMB # 1850-New v.1

Submitted by
National Center for Education Statistics
U.S. Department of Education

October 2016
revised December 2016

Contents

B. Collection of Information Employing Statistical Methods.....	1
1. Respondent Universe.....	1
a. Institution Universe.....	1
b. Student Universe.....	1
2. Statistical Methodology.....	1
3. Methods for Maximizing Response Rates.....	4
4. Tests of Procedures and Methods.....	4
5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study.....	10
C. References.....	10

Tables and Figures

Table 1.....Distribution of the <i>Financial Aid Nudges 2017</i> sample, by treatment group	2
---	---

Figure 1..... Power Curves
..... 8

B. Collection of Information Employing Statistical Methods

This submission requests clearance for the intervention methods and materials planned for the NCER-NPSAS grant study entitled, *Financial Aid Nudges 2017: A National Experiment to Increase Retention of Financial Aid and College Persistence*. An overview of the design is provided below; for more information see Part A.2. Similarly, materials are described briefly below and provided in appendix A.

1. Respondent Universe

The respondent universe for the *Financial Aid Nudges 2017* study is that of the 2015-16 National Postsecondary Student Aid Study (NPSAS:16). Details about the respondent universe for NPSAS:16, approved by OMB (#1850-0666 v.17-19), are provided below in sections 1a and b.

a. Institution Universe

To be eligible for NPSAS:16, an institution was required, during the 2015-16 academic year, to:

- Offer an educational program designed for persons who had completed secondary education;
- Offer at least one academic, occupational, or vocational program of study lasting at least 3 months or 300 clock hours;
- Offer courses that are open to more than the employees or members of the company or group (e.g., union) that administered the institution;
- Be located in the 50 states, the District of Columbia, or Puerto Rico;¹
- Be other than a U.S. Service Academy; and
- Have a signed Title IV participation agreement with the U.S. Department of Education.

Institutions providing only avocational, recreational, or remedial courses or only in-house courses for their own employees were excluded. The seven U.S. Service Academies were excluded because of their unique funding/tuition base.

b. Student Universe

The students eligible for inclusion in the NPSAS:16 sample were those enrolled in a NPSAS-eligible institution in any term or course of instruction between July 1, 2015 and April 30, 2016, who were:

- Enrolled in (a) an academic program; (b) at least one course for credit that could be applied toward fulfilling the requirements for an academic degree; (c) exclusively non-credit remedial coursework but who the institution has determined are eligible for Title IV aid; *or* (d) an occupational or vocational program that required at least 3 months or 300 clock hours of instruction to receive a degree, certificate, or other formal award;
- Not currently enrolled in high school; and
- Not enrolled *solely* in a GED or other high school completion program.

¹ Institutions in Puerto Rico were not eligible for NPSAS:12.

2. Statistical Methodology

The goal of the *Financial Aid Nudges 2017* study is to provide students with timely and relevant information regarding the importance of filing or refiling the Free Application for Federal Student Aid (FAFSA), the need to meet Satisfactory Academic Progress (SAP) requirements to retain aid eligibility, the importance of completing other financial aid forms and processes, the availability of additional financial resources and benefits available to students, and the benefit of making use of campus-based support resources. Participating students will be texted, using a texting platform developed by Signal Vine, and offered financial aid counseling by advisors employed by College Possible. The text message content is designed to inform students about the following topics:

- state- and institution-specific FAFSA priority refiling deadlines;
- how to file or refile the FAFSA, including how to prefill the form with responses from the prior year's FAFSA;
- SAP requirements;
- general campus-based resources students can access to refile the FAFSA, get questions answered about financial aid, and receive academic support to meet SAP requirements;
- financial aid award letters and how to interpret them;
- how to interpret and evaluate supplementary loan options for the following year;
- how to interpret and select tuition bill payment options;
- prompts to explore summer employment and internship options;
- guidance on course and major selection; and
- other forms of financial support such as EITC, food stamps, emergency aid, and emergency housing.

Matching to three administrative databases – the Central Processing System (CPS), the National Student Loan Database System (NSLDS), and the National Student Clearinghouse (NSC) – will allow the grantee and her research team to examine the impact of the information provided on FAFSA filing, financial aid receipt, and college persistence and success.

The student sample for the *Financial Aid Nudges 2017* study, shown in table 1, will be selected from among the NPSAS:16 student sample members who (1) completed the NPSAS:16 survey, (2) were not selected for participation in the 2016/17 Baccalaureate and Beyond Longitudinal Study (B&B:16/17; OMB# 1850-0926), (3) were in their first three years of their undergraduate education, and (4) agreed to participate in follow-up research in response to the following NPSAS:16 survey question:

We are almost done with this survey. But first, we want to let you know that some students may be invited to participate in follow-up studies to learn more about their education and employment experiences after completing this survey. These follow-up studies will be led by external researchers not affiliated with the Department of Education. We would like to seek your permission to allow RTI to re-contact you on behalf of one of the external researchers. Your participation in future studies is completely voluntary, but there is no substitute for your response. Are you willing to be contacted about these future studies? Yes/No

Table 1. Distribution of the *Financial Aid Nudges 2017* sample, by treatment group

Treatment group	Sample size
Total	8,500
Group 1 – Information and Nudges (with three sub-groups of different content variations)	3,400
Group 2 – Information, nudges, and assistance	3,400
Group 3 – Control	1,700

The three groups are defined as follows:

- **Group 1 - Information and nudges only.** These students will receive texts with information about FAFSA filing and refiling, and reminders to complete other stages of the financial aid process in advance of relevant state- and institution-specific priority filing deadlines. They will also receive information on SAP and other forms of financial support, as well as information about the other topics indicated above. All outgoing texts will be automated, and predetermined. Automated replies will be sent in response to key words included in texted responses from students. There will be three variations of content framing within Group 1, with Group 1 divided into three equal-sized subgroups. One content variation will be a basic presentation of information, and will be sent to Group 1A. The second variation will leverage positive social norms/pressure to encourage students to complete important tasks, and will be sent to Group 1B. For instance, one message will say :“(1/2) Hi, it’s Lara. Just checking in about FAFSA. About 70 percent of college students have completed their FAFSA by this time of year.” “(2/2) Have you finished FAFSA yet? Reply YES or NO.” A third variation will provide students with concrete planning prompts to encourage them to schedule specific days and times to complete important tasks (commitment device or reminder system), and will be sent to Group 1C. For instance, a message will say: “Hi it’s Lara again! I want to make sure you get financial aid to help pay for next year in college. Did you finish FAFSA in the fall? Reply YES or NO. If NO: Can you pick a day by next week when you can get this step completed? (Reply Monday, Wednesday or Friday or CAN’T).”
- **Group 2 - Information, nudges, and assistance.** In addition to the content similar to that received by students in Group 1, these students will receive explicit prompts to text back to interact with a financial aid advisor.
- **Group 3 - Control.** These students will not receive any text messaging outreach. They will still be receiving any financial aid and other college-success information and supports that are typically offered by their postsecondary institutions and other relevant agencies, such as the U.S. Department of Education.

A randomized controlled trial will be conducted to obtain unbiased causal estimates of the impact of the text messaging outreach on FAFSA refiling, the size and structure of students’ financial aid packages, maintenance of SAP, and, ultimately, college persistence and success. RTI will conduct the randomization. A sample of approximately five students per campus across 1,700 postsecondary institutions is anticipated. From the anticipated sample of approximately 8,500 students, students will be distributed across the three experimental conditions with approximately 3,400 students assigned to each, Group 1 and Group 2, and 1,700 students assigned to the control condition, Group 3. Greater shares of students can be assigned to the two treatment conditions given the expectation that differences between the treatments

may be smaller than differences between the control condition and each of the treatments. After Group 1 students are selected, they will be further subdivided into three evenly sized subgroups corresponding to the three different variations of the information and nudge outreach without text-based assistance. In the process of conducting the randomization, we will prioritize ensuring balance in our sample by institution type and also by whether a student is attending college full-time or part-time at the time of the NPSAS data collection. Therefore, we will stratify our sample into four groups – (1) full-time students at two-year institutions; (2) part-time students at two-year institutions; (3) full-time students at four-year institutions; and (4) part-time students at four-year institutions, and we will conduct our randomization within these strata. RTI will check to assure that the students are balanced across experimental groups on a host of covariates.

This structure for the randomization prioritizes the ability to detect differences between outcomes for Group 2 and outcomes pooled across the subgroups for Group 1, as well as comparing each of these groups to the control condition. This approach also achieves adequate power for exploring possible differences among the Group 1 subgroups.

The student-level randomization described above will allow maximization of statistical power for detecting treatment effects of interest. Because the study sample will include only a small and randomly-selected share of students from any particular NPSAS-participating campus, there is minimal risk of contamination across participating students on the same campus. Given that students opted to participate in follow-up research to NPSAS:16, low rates of opt out are anticipated. Students who opt out of receiving text-based outreach, students who transfer to a new postsecondary institution, and students who exit postsecondary education altogether will be maintained in the analytic sample for *Financial Aid Nudges 2017*. The data sources that will be utilized for tracking outcomes are not institution-specific, so nearly full coverage on outcomes is anticipated for students in the sample. No analytic issues are anticipated due to potential sample attrition.

3. Methods for Maximizing Response Rates

The success of the *Financial Aid Nudges 2017* intervention relies on the premise that text messaging is an effective strategy for communicating with students. Consenting students provided a cell phone number at the time of the NPSAS:16 interview. Nonetheless, the research team’s ability to implement the intervention requires that these numbers remain active during the duration of the intervention. RTI will track the number of nonworking cell phone numbers and will use data from Signal Vine to obtain information on when cell phones become inactive during the intervention and when students opt out of receiving messages after the intervention has commenced.²

Messages to Group 1 (information and nudges), which will include text strings, will be automated using a web-based platform provided by subcontractor Signal Vine. Students may respond to the automated texts, but responses to their texts will be pre-scripted and sent automatically in response to a student’s use of key word triggers. The content variations described above are designed to test different behaviorally-

² In previous studies, approximately 4% of students opted out (Castleman, B. L., & Page, L. C. (2015, 2016)).

informed strategies for maximizing student engagement with and responsiveness to the messages sent.

Text messages to Group 2 (information, nudges, and assistance) will include similar automated outreach as Group 1, but will also contain explicit prompts to students to write back if they have questions or need assistance. More specifically, each Group 2 message will be comprised of two core components. The first core component will align directly with the topical focus of the Group 1 message. For instance, if the Group 1 message is about FAFSA filing, the first component of the Group 2 message will be information about FAFSA filing. If the Group 2 message is about meeting SAP, the first component of the Group 2 message will contain information about SAP, and so on. The second component of the Group 2 text will be an explicit invitation for the student to write back with a question or to connect with an advisor if they need assistance. The research team has drafted these prompts with a maximal conversational tone which the team has found, in prior research, to elicit high response and engagement rates (Avery et al., in progress; Barr et al., in progress). Examples of these invitations are included in Appendix A. Students who respond to the text messages will receive, within minutes, remote advising assistance from College Possible (advisors will be available during afternoons, nights, and weekends, when students are most likely to seek assistance). College Possible staff will be able to use additional interactive technologies, such as screen sharing and document collaboration, to provide more in-depth advising assistance to students if necessary (e.g., review a financial aid award letter the student received).

Customized segments will also be incorporated into the text messaging schedule for students who do not engage with the first several messages. In Group 1, students will be prompted to respond with a keyword (e.g. "FAFSA" or "NEXT") to receive messages with extended content about the particular topic on which that message was focused (e.g., FAFSA filing, meeting SAP, making use of campus resources). For students who do not respond to any of the keyword prompts in the first several messages, the "auto-flow" features of the text messaging system will be used to send them a short series of messages that do not require a keyword prompt. For instance:

In case you haven't started the FAFSA yet, we wanted to share with you a few steps you can take to get started.

AUTO FLOW TO:

(1/4)Remember you'll need your FSA ID to digitally sign your application at fafsa.gov. Text "NEXT."

AUTO FLOW TO: (2/4) Check your email regularly to see if you get any emails about FAFSA. Contact the aid office if you are required to verify your income info. Text "NEXT."

Note that the "auto flow" of messages is still limited to 2-3 messages so as to not overwhelm the student with a long series of messages that they have not actively indicated they wish to receive.

In Group 2, if students have not responded to the invitation to connect with an advisor in the first several messages, they will be sent a message asking if they would prefer to receive automated information about important steps they can take to maximize

financial aid or meet SAP, and provide keyword options, like Group 1 receives, to receive more content.

For the customized segments for non-responders, the research team may incorporate a Sequential Multiple Assignment Randomized Trial (SMART) experimental design into this segmentation. SMART designs allow researchers to understand how adaptations to intervention design during a trial affect individual outcomes by pre-specifying assignment of participants to different pathways based on some value of intervention response or participation (in this case, not responding to one of the first several text messages (Lei, et al., 2012)). A SMART design would allow the research team to investigate whether their modifications for non-responders are more effective at engaging students who might not have otherwise interacted with the system.

In addition, in order to maximize response rates, focus group(s) will be conducted in December 2016 (OMB# 1850-0803 v.183) to test whether the language of the text messages is understandable to a group of low-income and first generation students. If the focus group(s) suggest changes should be made to text message language, those revisions will be submitted in December 2016 as part of this submission and/or via a change memo to OMB in early 2017.

Signal Vine will track the extent to which students responded to messages to request assistance, and participating College Possible advisors will track their interactions with students in formal logs. At the end of the 6-month intervention, matched de-identified data from the advisor logs will be compared with the text message providers' records of requests for assistance from students. This comparison will allow an assessment of several important measures of program implementation and fidelity, including the proportion of students who requested help, the proportion to whom advisors reached out and with whom advisors successfully interacted, the average time between recipients' requests for help and advisors' outreach, and the extent to which the help advisors provided aligned with the message topic to which recipients responded.

4. Tests of Procedures and Methods

During the period of active intervention, students will be immersed in the everyday routines of college students: taking classes, participating in extracurricular activities, working, spending time with friends and family, and commuting. For some students, the messages texted as part of the grant study interventions will arrive while they are relaxing in their dorm; for others, messages will arrive while they are at home with family or at a job. Furthermore, students will receive messages from an organization with which they do not have a prior relationship. The student responses, interactions with advising staff, behaviors, and outcomes that result from the intervention will, therefore, provide an informative and authentic view on what might be expected if, for instance, Federal Student Aid were to undertake a texting campaign to promote FAFSA filing and refiling and compliance with SAP requirements. Critically, the grantee will be estimating the effects of *sending the texts* (the Intent-to-Treat) rather than the effects of *receiving the texts*, since it is only possible for governments and schools to ensure the former and not the latter.

Process outcomes. To examine whether students in the treatment groups react to the interventions, data will be collected on a variety of intermediate process outcomes, including:

- measures of text response and advisor interaction,
- whether students click on the links provided for campus financial aid and academic support resources,
- whether students click on the links for additional information, and
- the number of unique views of each provided video.

Financial aid. Using the NSLDS and CPS, the following will be examined:

- rates and dates of student FAFSA refiling,
- the intervention’s impact on students’ total aid packages, and
- the composition of those aid packages to investigate whether grants constitute a greater share of their financial aid package rather than loans, which increase students’ postsecondary debt burden.

Retention and completion. The ultimate goal of the interventions is to increase both student retention and completion. The NSC makes it possible to examine whether the interventions impact student retention—both term-to-term and year-to-year – and student completion outcomes using the NSC’s Degree Verify data. For students at 2-year institutions, and for all sample students (those in their first 3 years of college), the grantee will be able to observe whether the intervention impacts completion in 100%, 150%, and 200% of normal time. Given the limits of the project timeline, for students at 4-year institutions, the grantee will observe completion in 100% and 150% of normal time for students in their second and third years of college at the time of the intervention. For students beginning college at the time of the intervention, the grantee will observe completion in 100% of normal time.

Information collected during the NPSAS:16 survey, including on the institution attended, student year in college, and Pell eligibility will be utilized to examine moderators of treatment impacts. The grantee hypothesizes that two key mediators for effects on college retention could be (a) filing of the FAFSA and (b) receipt of financial aid. These will be measured in the ways explained above. In addition, mediators will be examined related to how the treatment is received directly from Signal Vine and College Possible (e.g., whether students access the information provided and use personalized assistance for help in refiling).

Rigorous causal methods will be used to determine the success of the interventions and to investigate a set of targeted research questions in this study as described below.

1. Does offering students simplified information about the requirements for retaining financial aid and prompts delivered via text messaging “nudges” increase FAFSA refiling rates and/or SAP compliance rates?
 - a. Is text messaging more effective if the texts include an offer to the student to interact with an advisor? Is that variation of nudging also more cost-effective?

- b. Is text content that leverages positive social norms or that provides planning prompts more effective at eliciting student response and engagement than basic information?
 - c. Does the impact of text-messaging vary by institutional sector (e.g., 2-year versus 4-year colleges and universities)? Does it vary by a student's Pell-eligibility or year in school?
 2. To what extent does text-messaging increase the amount of financial aid that students receive? How does it impact the composition of their financial aid packages (e.g., the mix of grants and loans)?
 3. To what extent does text-messaging increase the rates of retention to the next year of college and rates of completion?

The grantee hypothesizes that the components of the interventions described above for Groups 1 and 2 will lead to several behavioral changes among students. By conveying simplified information about FAFSA filing and refiling, along with information about state- and institution-specific priority filing deadlines, students' awareness of the need to file/refile and motivation to file/refile will be increased. The financial benefits associated with filing/refiling will be conveyed in advance of these deadlines and keep FAFSA at the top of students' minds as a result of the ongoing prompts. The grantee anticipates observing an overall increase in the share of students who file/refile their FAFSA and earlier FAFSA filing and refiling, when compared to students who do not receive these prompts.

It is further anticipated that the impacts of the text messages on FAFSA filing/refiling and the timing will be more pronounced among students in the "information, nudges, and assistance" group, and that by offering these students individualized, real-time assistance with the FAFSA, students' tendency to procrastinate in the face of onerous hassles will be reduced as advisors help them work through challenges or confusion associated with the process. Given NPSAS:16's existing ability to draw student-level data from CPS and NSLDS, the goal is to leverage these same data sources to generate measures of whether the interventions do in fact lead to earlier and higher overall rates of FAFSA filing and refiling. It is expected that the effect of the texts will be moderated by both student- and institution-level characteristics. For instance, the effects of the proposed treatments are likely to be greater for students with less access to financial aid information and support for completing the FAFSA, such as students who are the first in their family to go to college. It is also anticipated that the texts will be more effective for students who, based on the demands for their time, are less likely to devote attention to complete the FAFSA, such as students who are working a substantial number of hours or students who have dependents. These student-level moderators will be available from the NPSAS:16 data. Institution-level moderators may include characteristics such as the ratio between financial aid advisors and students, or the share of the student population that commutes to campus. While some of these institution-moderators will be available from the Integrated Postsecondary Educational Data System (IPEDS), the grantee may also have to rely on proxies, such as the instructional dollars each institution expends per student.

The grant study is designed to isolate the causal impact of two important mediators—content framing and the offer of individualized financial aid advising—for how the text messages affect FAFSA filing/refiling, meeting SAP, and subsequent educational outcomes. Because Group 2 (“information, nudges, and assistance”) will be randomly assigned to receive invitations to write back to engage with an advisor if they have questions or need assistance, while Group 1 (“information and nudges”) will not, the grantee will be able to identify the effect of offering financial aid advising on whether and when students file or refile the FAFSA, and on whether they persist in college. Because students within Group 1 will be randomized to receive different content frames, the grantee will also be able to identify the effect of positive social norms or planning prompts relative to basic information on whether and when students file or refile the FAFSA, and on whether they persist in college.

Due to limited statistical power, the grantee will not be able to experimentally assess the impact of other mediators, such as whether the texts are effective because they provide students with information they did not already have, or because they prompt students to invest in a task that they had been putting off. The grantee will, however, be able to rely on a variety of descriptive measures to explore these mediators. For instance, click rates on links included in the messages will be recorded, and the frequency with which embedded videos are watched, to infer whether the texts are providing new information to students.

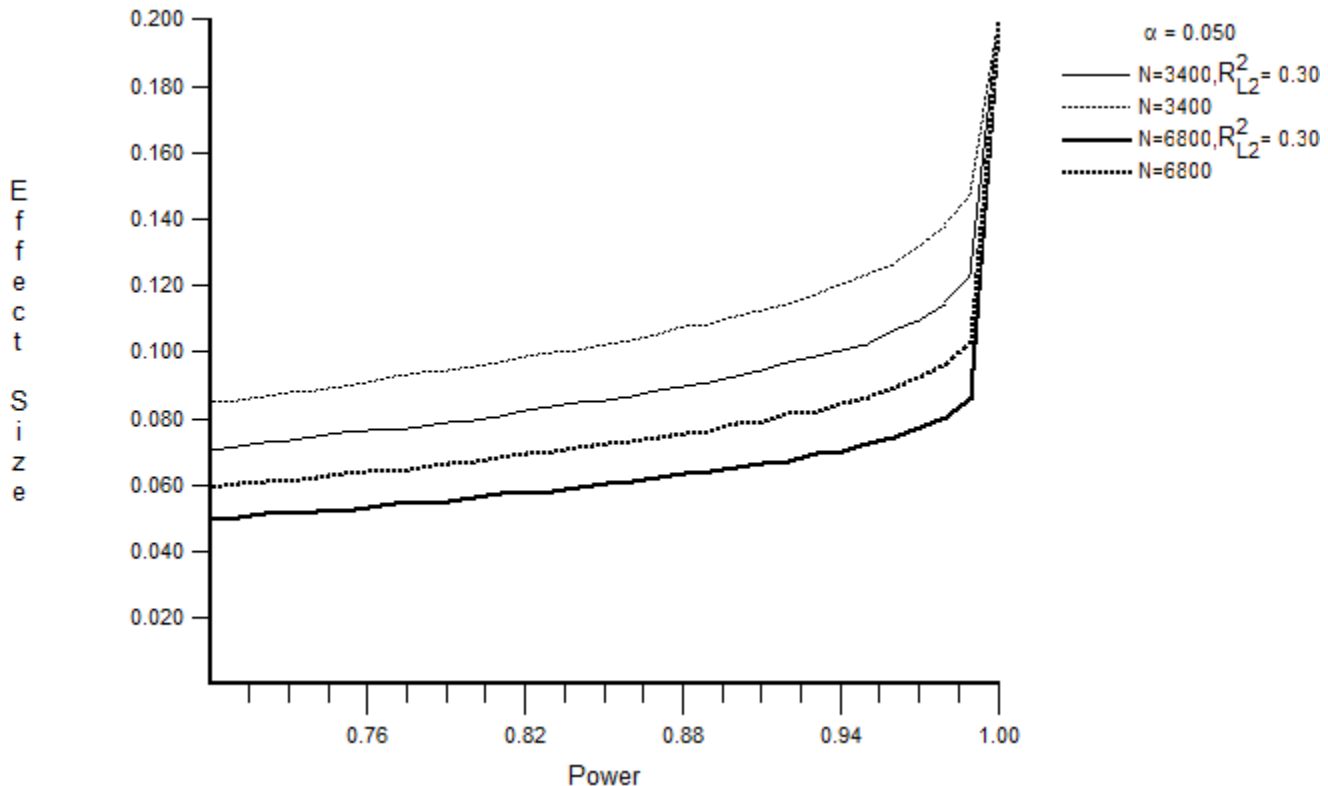
The grantee will also analyze the content of interactions between students and advisors in the “information, nudges, and assistance” Group 2 to examine the types of questions students ask and the areas for which they seek help. Leveraging data from CPS, the grantee will compare the timing of when students start, submit, and complete the FAFSA across the two treatment groups, relative to when the grantee sends FAFSA-filing related reminders, to see if students in the treatment groups are more likely than students in the control group to begin or continue working on FAFSAs in the days following when the prompts are delivered. While the frequency with which outgoing messages are sent will be controlled, students’ decisions about how they engage with the text content (e.g., by watching videos; following automated, interactive messaging modules) are likely endogenous to other student factors affecting FAFSA filing/refiling and subsequent outcomes. The grantee can nonetheless descriptively explore the relationship between engagement with the text content and students’ subsequent educational outcomes, controlling for the rich set of baseline measures to which the grantee will have access.

As noted, the sample will include approximately two students per primary experimental condition (e.g., Group 1 or Group 2) and one student in the control per campus, across 1,700 campuses. Power calculations associated with this experimental structure are presented below. In the power calculations, the focus is on the contrast between the control condition and a single treatment condition. In this way the calculations can be considered conservative, given that the full sample will be larger. After presenting power calculations for this particular contrast, analogous calculations for other contrasts of interest are discussed.

Optimal Design Plus Empirical Evidence (Version 3.0) (Raudenbush et al., 2011) will be utilized to conduct the power analyses. Power is calculated assuming a simple

student-level randomization. A 0.05 level of statistical significance is utilized and the control group rate of success with binary outcomes of interest is set at 0.50.³ Assuming a blocking variable (institutional sector) that explains 10% of the variation in the outcome and student-level characteristics that explain 20% of the variation in any given outcome, quite modest impacts of the intervention will be detectable, given the ability to randomize students at the student level. Specifically, with these parameters, an 80% power to detect effects of each treatment arm (Group 1 or Group 2) compared to the control (Group 3) is achieved, on the order of 0.081 to 0.096 standard deviations, as illustrated in Figure 1.

Figure 1. Power Curves



If binary outcomes of interest (e.g., FAFSA filing/refiling, next-year persistence) with baseline control group rates of 50% are assumed, this implies minimal detectable effects on the order of 4.1 to 4.8 percentage points.⁴ With two observations per active treatment group and with a similar set of parameters, there will be power to detect differences between Group 1 and Group 2, the treatment arms, on the order of 2.9 to 3.4 percentage points (see Figure 1).

Finally, given the further sub-dividing of the Group 1 students described above, approximately 1,133 students will be assigned to each of the three different message types. Again assuming that covariates explain 30% of the variation in the outcomes, the research team estimates that the model will achieve 80% power to detect differences on the order of 0.10 standard deviations, or approximately 5 percentage points.

These power calculations are considered to be conservative in nature, as they do not

³ With a binary outcome, power will be increased the further the control group rate is from 0.50. Therefore, a control group rate of 0.50 is utilized for power calculations, as it is the most conservative choice possible.

⁴ Assuming a control group enrollment rate of .50, the standard deviation is $\sqrt{.5 \times .5} = .5$

take into account additional precision to be achieved by estimating treatment effects simultaneously. Because of the conservative nature of these calculations, the effect sizes reported here serve as an upper bound on the size of the detectable effects. Nevertheless, given the research team's prior experience with interventions such as these, impacts of this magnitude are judged to be entirely plausible.

In sum, there is power to detect quite small impacts, and the design and sample sizes will allow estimation of the impact of the interventions with a high degree of precision.

Given the use of a randomized control trial design, straightforward regression models will be utilized to answer the research questions related to intervention impact. For example, we will utilize the following general model to investigate the primary impacts of the intervention on both groups 1 (overall) and 2:

$$Y_{ij} = \beta_0 + \beta_1 \text{TREAT1}_i + \beta_2 \text{TREAT2}_i + \beta_3 X_i + \beta_4 Z_j + \varepsilon_i, (1)$$

where for student i and postsecondary institution j , TREAT1 is a binary indicator for assignment to treatment Group 1, and TREAT2 is a binary indicator for assignment to treatment Group 2.⁵ Y_{ij} represents the outcomes of interest, such as FAFSA refiling, under investigation. X represents a vector of student-level baseline covariates anticipated to be incorporated into the model, including demographic, socioeconomic, and academic baseline measures, Z represents a vector of institution-level baseline covariates such as institution type and sector, and ε_i is the student-level error term.

In (1), β_1 and β_2 are the primary parameters of interest. β_1 represents the improvement in a given outcome of interest over the control group rate that is induced by treatment 1. β_2 represents the analogous improvement that is induced by treatment 2. Estimates of these parameters that are positive and both practically and statistically significant will be interpreted to indicate that these treatments led to improvements in key college-success outcomes. A post-hoc linear hypothesis test will be used to investigate whether impacts of one treatment are significantly larger than impacts of the other (e.g., treatment 1). If, for example, the Group 2 treatment impacts are larger than the Group 1 treatment impacts, it will be concluded that the opportunity to connect with an advisor one-on-one led to meaningful improvements in postsecondary outcome measures, above and beyond one-way outreach to provide information and reminders. This model will be amended in order to assess variation in impacts across the group 1 variations. This will be achieved by amending model (1) to incorporate separate treatment effects for the three group 1 subgroups in place of the overall group 1 treatment effect indicated above.

Based on prior work from members of the grantee's research team, it is plausible that the impact of the interventions may vary by institution characteristics and student-level characteristics. For example, prior work focused on FAFSA renewal was particularly effective within community college settings (Castleman & Page, 2016). In addition, it is reasonable to expect for FAFSA filing and renewal to be most critical for students from the lowest-income families, for whom failure to retain financial aid could represent a particular barrier to postsecondary continuation. To investigate whether either of the interventions has a differential impact by salient student- or institution-

⁵ Note that this is an intent-to-treat model based on assignment to treatment. Critically, the effects of *sending the texts* (the Intent-to-Treat) rather than the effects of *receiving the texts* will be estimated, given that it is only possible for governments and schools to ensure the former and not the latter.

level characteristics, equation (1) will be modified by interacting the relevant treatment indicators with student-level measures (such as a measure of family wealth) and institution-level measures (such as institution type). Positive and significant coefficients on these terms will indicate that the treatment is more impactful for certain subsets of students and/or in certain types of postsecondary settings.

Cost-benefit analysis requires several components: (1) the effects of the treatment on medium- and long-run outcomes such as graduation; (2) an estimate of the relationship between degree receipt and subsequent earnings, and (3) estimates of the cost per student of implementing the two different treatments. Estimation of (1) and (3) is described above. β_1 and β_2 are the treatment effects of the two interventions on persistence (or, if possible, college graduation). The cost of providing the two treatments will be calculated using the expenses incurred for Signal Vine to do the texting and the professional advisers for the personalized assistance.

These cost estimates and treatment effects will enable the calculation of a cost per additional student retained in college and a cost per additional graduation. This will permit a comparison of the relative cost effectiveness of the two interventions and enable a comparison to other interventions for which a cost has already been calculated per additional student induced into college (see Carrell & Sacerdote, 2013, for a discussion and examples). To estimate (2), a rudimentary cost benefit analysis can also be conducted using existing estimates of private and social returns to an additional year of college. For example Gunderson and Oreopoulos' (2010) survey of the literature suggests that an additional year of college provides benefits of \$5,000 per worker, and this midpoint estimate can be compared to the cost per additional student retained for another year of college.

5. Reviewing Statisticians and Individuals Responsible for Designing and Conducting the Study

The *Financial Aid Nudges 2017* NCER-NPSAS grant study requires substantial coordination between the grantee, RTI, and NCES. The grantee's research team is responsible for research design, data analysis, and dissemination of results, and includes the following individuals: Dr. Sara Goldrick-Rab, Dr. Ben Castleman, Dr. Lindsay Page, Dr. Jed Richardson, and Dr. Bruce Sacerdote. The following staff members at RTI are responsible for sample selection, respondent contacting and follow-up, data collection and processing, and weighting of the data (if applicable): Ms. Kristin Dudley, Mr. Jeff Franklin, Mr. Peter Siegel, and Dr. Jennifer Wine. Dr. Tracy Hunt-White and Dr. Sean Simone, from NCES, and Dr. James Benson, from NCER, are the statisticians responsible for ensuring that the confidentiality of NPSAS:16 sample members is protected (including ensuring that proper data security protocols are in place and that NCES Statistical Standards are met) and for the general oversight of the NCER-NPSAS grant program.

C. References

- Ashraf, Nava, Dean Karlan, and Wesley Yin. Tying Odysseus to the mast: Evidence from a commitment savings product in the Philippines. *The Quarterly Journal of Economics* 121.2 (2006): 635-672.
- Babcock, Philip, and John Hartman. Networks and workouts: Treatment size and status specific peer effects in a randomized field experiment. NBER Working Paper No. 16581. (2010)
- Bettinger, Eric P., et al. The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *The Quarterly Journal of Economics* 127.3 (2012): 1205-1242.
- Bird, Kelli, and Ben Castleman. Here today, gone tomorrow? Investigating rates and patterns of financial aid renewal among college freshmen. *Research in Higher Education* 57.4 (2016): 395-422.
- Broton, Katharine M., and Sara Goldrick-Rab. Public Testimony on Hunger in Higher Education Submitted to the National Commission on Hunger. (2015).
- Burstyn, Leonardo, and Robert Jensen. How Does Peer Pressure Affect Educational Investments? NBER Working Paper No. w20714. (2014)
- Cannon, Russell, and Sara Goldrick-Rab. (2015). Why didn't you say so? Experimental Impacts of a Financial Aid Call Center. Preliminary working paper.
- Carrell, S. E., & Sacerdote, B. (2013). *Why do college going interventions work?* (No. w19031). National Bureau of Economic Research.
- Castleman, Ben, and Lindsay Page. (2016). Freshman year financial aid nudges: An experiment to increase financial aid renewal and sophomore year persistence. *Journal of Human Resources* 51.2 (2016): 389-415.
- Castleman, Ben, and Lindsay Page. (2015a). Beyond FAFSA completion. Change. January-February 2015. Retrieved from <http://www.changemag.org/Archives/Back%20Issues/2015/January-February%202015/beyond-fafsa-full.html>.
- Castleman, Ben, and Lindsay Page (2015b). Summer nudging: Can personalized text messages and peer mentor outreach increase college going among low-income high school graduates? *Journal of Economic Behavior and Organization* 115 (July 2015): 144-160.
- Deming, David, and Susan Dynarski. Into college, out of poverty? Policies to increase the postsecondary attainment of the poor. NBER Working Paper No. w15387. (2009)
- Duke-Benfield, Amy Ellen, and Katherine Saunders. Benefits access for college completion. Center for Postsecondary and Economic Success at CLASP. (2016).
- Dynarski, Susan, and Mark Wiederspan. Student aid simplification: Looking back and looking ahead. *National Tax Journal* 65.1 (2012): 211-234.
- Goldrick-Rab, Sara, Kelchen, Robert, Harris, Douglas N., and James Benson. Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion. *American Journal of Sociology* 121.6 (2016): 1762-1817.

- Gunderson, M., & Oreopoulos, P. (2010). Returns to education in developed countries.
- Kelly, Andrew, and Sara Goldrick-Rab. Reinventing financial aid: Charting a new course to college affordability. Cambridge, MA: Harvard Education Press. (2014)
- Laibson, David. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics* 112.2 (1997): 443-477.
- Lei, H., Nahum-Shani, I., Lynch, K., Oslin, D., & Murphy, S. A. (2012). A “SMART” Design for Building Individualized Treatment Sequences. *Annual Review of Clinical Psychology*, 8, 10.1146/annurev-clinpsy-032511-143152. <http://doi.org/10.1146/annurev-clinpsy-032511-143152>
- McDonnell, Rachel Pleasants, and Lisa Soricone. Promoting persistence through comprehensive student supports. *Jobs For the Future* (2014).
- McKinney, Lyle, and Heather Novak. FAFSA filing among first-year college students: Who files on time, who doesn't, and why does it matter? *Research in Higher Education* 56.1 (2015): 1-28.
- Novak, Heather, and Lyle McKinney, L. The consequences of leaving money on the table: Examining persistence among students who do not file a FAFSA. *Journal of Student Financial Aid* 41.3 (2011): 5-23.
- St. John, Edward, Hu, Shouping, and Tina Tuttle. Persistence by undergraduates in an urban public university: Understanding the effects of financial aid. *Journal of Student Financial Aid* 30.2 (2000): 23-37.
- Raudenbush, S. W., Spybrook, J., Congdon, R., Liu, X. F., Martinez, A., & Bloom, H. (2011). Optimal design software for multi-level and longitudinal research (Version 3.01)[Software]. Available from www.wtgrantfoundation.org.
- Schudde, Lauren, and Judith Scott-Clayton. Pell grants as performance-based aid? An examination of satisfactory academic progress requirements in the nation's largest need-based aid program. A CAPSEE Working Paper. New York, NY: Center for Analysis of Postsecondary Education and Employment. (2014)
- Thaler, Richard, and Cass Sunstein. *Nudge: Improving decisions about health, wealth, and happiness*. New York, NY: Penguin Books. (2009)
- Wisconsin HOPE Lab. (2015). What we're learning: Satisfactory Academic Progress. Wisconsin HOPE Lab Data Brief 15-01. Available at <http://wihopelab.com/publications/>