margaret.cahalan@pellinstitute.org

---

From:margaret.cahalan@pellinstitute.org
Sent:7 Apr 2014 11:16:01 -0400
To:info@whatworks.ed.gov
Subject:IES WWC Website: Contact Us: Other, Reference ID Number: 2086160509

info@whatworks.ed.gov, this email was automatically sent through the Contact link on the WWC website.

From: margaret.cahalan@pellinstitute.org

Message: David Goodwin and I, Margaret Cahalan have prepared a detailed formal request for a Quality Review of WWC ratings of the 2004 and 2009 Mathematica reports from the National Evaluation of Upward Bound.

I am requesting the email address to which I should attach the material pertaining to this case.

# Attachment A: *Excerpt of UB Study Conclusions from the Executive Summary of the 2009 Mathematica Final Report*

**Executive Summary conclusions excerpted from the report:** **Seftor, N. S., Mamun, A., & Schirm, A. (2009).** ***The impacts of regular Upward Bound on postsecondary outcomes 7–9 years after scheduled high school graduation*****. Princeton, NJ: Mathematica Policy Research.**

"By comparing the study's treatment group to its control group, this evaluation estimates the value-added effect of the opportunity to participate in Upward Bound—an unusually intensive precollege program—for the students who seek that opportunity and are eligible to participate in the program. The main finding are:

- ***Upward Bound had no detectable effect on the rate of overall postsecondary enrollment or the type or selectivity of postsecondary institution attended for the average eligible applicant.*** *About four-fifths of both treatment group members and control group members attended some type of postsecondary institution, including four-year institutions, two-year colleges, and vocational schools, and the estimated impact is an increase of less than 2 percentage points in the rate of enrollment (effect size = 4 percent). For enrollment at four-year colleges and universities, the estimated impact is 1 percentage point (effect size = 3 percent). These effects are not statistically significant.*
- ***Upward Bound had no detectable effect on the likelihood of applying for financial aid, or, the likelihood of receiving a Pell Grant.*** *The 1 and 2 percentage point increases in the rates of financial aid application and Pell Grant receipt (effect sizes = 3 and 5 percent) are not statistically significant.*
- ***Upward Bound increased the likelihood of earning a postsecondary certificate or license from a vocational school. It had no detectable effect on the likelihood of earning a bachelor's degree or the likelihood of earning an associate degree.*** *While about 4 percent of control group members received a vocational certificate or license, nearly 9 percent of treatment group members did, implying an impact of 5 percentage points (effect size = 23 percent). The impacts on receiving any postsecondary credential and receiving a bachelor's degree are 2 and 0 percentage points (effect size = 5 and 0 percent), respectively, and are not statistically significant.*
- ***Upward Bound increased postsecondary enrollment or completion rates for some subgroups of students.*** *For the subgroup of students with lower educational expectations at baseline—that is, the students who did not expect to complete a bachelor's degree—Upward Bound increased the rate of postsecondary enrollment and the likelihood of receiving a degree, license, or certificate by 6 and 12*

*percentage points, respectively, raising the overall postsecondary completion rate to about the level observed for students with higher expectations. Because targeting on the basis of lower educational expectations might be challenging if it creates an incentive for applicants to understate their expectations, further analyses were conducted to examine the effects of Upward Bound on subgroups that might be more readily targeted. According to these exploratory analyses, Upward Bound increased postsecondary enrollment rates for students who were in tenth grade or above at the time of application, students who took a mathematics course below algebra in ninth grade, and students with a ninth grade GPA above 2.5. The estimated impacts were 3, 7, and 3 percentage points, respectively. Additional analyses suggest that Upward Bound also had positive impacts on postsecondary outcomes for some other subgroups defined by student- and project-level characteristics.*

- ***Longer participation in Upward Bound was associated with higher rates of postsecondary enrollment and completion.*** *An additional year of Upward Bound participation was associated with a 9 percentage point increase in the rate of enrollment at four-year institutions and a 5 percentage point increase in the likelihood of receiving a bachelor's degree. Completing the Upward Bound program was associated with increases of 27 and 21 percentage points, respectively. These findings are based on nonexperimental methods, and the validity of causal inferences based on these estimates depends on the validity of strong assumptions. "*

# Attachment B: Documentation of Key Standards Violations in the Mathematica Reports from the National Evaluation of Upward Bound

**Submitted in Support of the Request for Rescinding of the WWC Rating of the Mathematica Reports as "Meets Evidence Standards, Without Reservations"**

**by Margaret Cahalan and David Goodwin**

To document the summary material presented in our letter, below we present a brief background concerning the unusual context of this issue followed by a detailed description of the major research standards violations manifest in the Mathematica reports. Attachment C provides additional documentation.

## How Did PPSS Staff First Become Concerned--Context and Background?

PPSS concerns first began in 2005 when the Mathematica lead analyst for the fourth follow-up of the long running study sent PPSS tabulations that showed that the "no impact" results were sensitive to only one of the 67 sampled projects.[1] This staff person revealed for the first time to ED that this one project (known as project 69) was carrying fully 26 percent of the total student weights and this meant that students from project 69 had weights that were 40 times those of the lowest project weighted sample members. The staff person was concerned that this project for unknown reasons had large significant negative impacts on college outcomes and given the sample members large weights the results and conclusions of "no impact" for Upward Bound were being driven by this project. Without this project UB clearly demonstrated statistically significant and educationally meaningful positive impacts in the analyses. PPSS staff asked the Mathematica UB project staff to investigate the issue further both in terms of the substantive reasons for the large negative impacts in project 69 and in terms of the unequal weighting issues. PPSS wished to understand what project 69 was doing wrong such that the project was seemingly decreasing rather than fostering college attendance and degree attainment. Given that this project was supposedly representative of the largest number of 4-year BA and above degree granting UB projects—these results could not be ignored. PPSS also wished to understand the role that the uneven weighting issue was playing in the estimates, having observed that the sub-group estimates were unstable and sensitive to issues of differential survey non-response and coverage issues. A few months later, without having gotten a response from Mathematica concerning these issues, the lead analyst working on the fourth follow-up report, left Mathematica. The fourth follow-up report was never revised to address PPSS internal reviewers concerns or put into review for publication[2] .

---

[1] Details on memo's and emails concerning PPSS concerns can be found in the COE Request for Correction, submitted to ED in 2012 and available on the COE website *The Council for Opportunity in Education (COE) Request for Correction* is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf

[2] The PPSS review of the fourth follow-up draft raised a number of issues some of which were similar to those of concern raised to the fifth follow up report. However, Mathematica reported they did not have resources to revise the fourth follow up report and also prepare the fifth follow up report. PPSS agreed to Mathematica's concentration on completing the fifth follow up report with remaining contract funds.

## The PPSS Quality Assurance (QA) Review

Given the concerns raised, in 2006, ED-PPSS began an internal QA review of the study. The QA review PPSS conducted involved an internal review and analysis of all data files from the study, as well as consultation and replication of PPSS re-analysis results by external statistical experts. Over the period of 2006-2007 as they became available, PPSS obtained from Mathematica copies of the data files through fifth follow-up and had them matched with the federal aid files. Data files reviewed included: the initial sampling frame, the baseline survey, 5 follow-up surveys, student transcripts, 10 years of federal aid files and National Student Clearinghouse (NSC) data.

PPSS also began to seek external statistical advice including consultation with Dr. James Chromy of RTI, the PPSS Technical and Statistical Assistance contractor at the time. Dr. Chromy is a well-respected Fellow of the American Statistical Society who has consulted on sampling issues for NAEP, NPSAS, BPS, B&B in the education field and also has expertise in experimental design issues in health clinical trials. Dr. Chromy and his statistical team gave PPSS advice concerning the sample design issues and also after obtaining the data files from the study, replicated the significant positive impact estimates that PPSS had found internally for the entire sample including Project 69, following the merge with the federal aid files using data files through the fourth follow-up. Unfortunately PPSS was not able to obtain a copy of the grantee population sampling frame from which the UB grantee sample was drawn from Mathematica until December of 2007, after the contract had ended and after the final report had been through an initial review.[3] After PPSS obtained the sampling frame and gained knowledge of the identity of sampled UB projects, PPSS became more concerned about the representational issues in addition to the unequal weighting and treatment and control group lack of balance.[4] The fact that the Project 69 was a former minority serving junior college and had an historical focus on certificates and service to CTE high schools was also helpful in understanding the large impacts on certificate attainment that had been observed. Gradually, over the period of 2006 to 2009 as study errors were identified, PPSS staff used NCES and WWC standards to mitigate these errors in a re-analysis described in more detail below.

## Major Standards Violations and Their Impact

As already noted in our letter, listed below are the key applicable sets of standards and guidelines used in identifying key flaws in the Mathematica Upward Bound reports.

- *U.S. Department of Education Information Quality Guidelines (ED Guidelines)*
- *Joint Committee on Standards for Educational Evaluation (JCSEE).* http://www.jcsee.org/)

[3] At the request of Dr. Chromy of RTI, the PPSS Technical and Analysis Support Contractor at the time, whom PPSS had asked to provide an independent review of the Mathematica UB sample design, PPSS requested the population sampling frame from Mathematica in March of 2007. However, due to delays in Mathematica's ability to locate the population frame it was 10 months before it was finally delivered to PPSS and after the contract had ended in December of 2007. Prior to the end of the contract in December of 2007, PPSS did not have knowledge of the identity of project 69.

[4] More recently, Dr. Chromy also reviewed the COE Request for Correction (2012) and gave advice concerning recommendations for mitigation. *The Council for Opportunity in Education (COE) Request for Correction submitted in 2012* is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf,

- *IES, National Center for Education Statistics (NCES) Statistical Standards*--- These may be accessed at the following site url: http://nces.ed.gov/statprog/
- *IES, What Works Clearinghouse Standards (WWC)* ---these may be accessed at the following site urls: http://ies.ed.gov/ncee/wwc/pdf/wwc_version1_standards.pdf
- http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1/
- *American Educational Research Association* (AERA) *Standards for Reporting on Empirical Social Science Research* in AERA Publications http://www.sagepub.com/upm-data/13127_Standards_from_AERA.pdf

Exhibits B-1 and B-2, (repeated from our letter) respectively first identify key specific applicable standards that have been violated and then summarize the 10 specific violations in the Mathematica reports. There follows a detailed discussion of each of the 10 violations with supporting exhibits documenting errors and re-analyses findings. Attachment C under a separate file in this package contains additional documentation and includes actual output from STATA tabulations taken from the 2009 report *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change? By Margaret Cahalan* a COE report published in 2009 and available at http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml

**Exhibit B-1**
**Key Information Quality Guidelines and Standards that are Applicable to the Concerns with Regard to the *Mathematica Upward Bound Report* s**

**Department of Education Quality Information Guidelines**

**Research and Evaluation information products should, at a minimum: …**

- Pose the research or evaluation question in a balanced and unbiased manner;
- Provide an unbiased test of the question; …
- Present conclusions that are strongly supported by the data; ….
- Confirm and document the reliability of the data, and acknowledge any shortcomings or explicit errors in any data that is included;
- The source of data should be reliable. The sample should be drawn from a complete list of items to be tested or evaluated, and the appropriate respondents should be identified, *correctly sampled,* and queried with survey instruments that have been properly developed and tested
- Appropriate steps should be taken to *ensure that the respondents are a representative sample*;

**What Works Clearinghouse Handbook of Procedures and Standards**

A study may fail to meet WWC evidence standards if ……..

- It does not include a valid or reliable outcome measure, or does not provide adequate information to determine whether it uses an outcome that is valid or reliable. …..
- The intervention and comparison groups are not shown to be equivalent at baseline
- The overall attrition and or differential attrition rate exceeds WWC standards for an area.
- The measures of effect cannot be attributed solely to the intervention

……

**NCES Statistical Standards Concerning Non-Response and Coverage**

- **STANDARD 2-2-4:** A nonresponse bias analysis is *required* at any stage of a data collection with a unit response rate less than 85 percent. The extent of the analysis must reflect the magnitude of the nonresponse (see Standard 4-4).
- **STANDARD 3-1-2:** NCES data collections that are used as sampling frames for other NCES surveys must strive for coverage rates in excess of 95 percent overall and for each major stratum. **STANDARD 3-1-3:** If there is not evidence of a coverage rate of at least 85 percent of the target population, then frame enhancements such as frame supplementation or dual frame estimation must be incorporated into the survey study design.

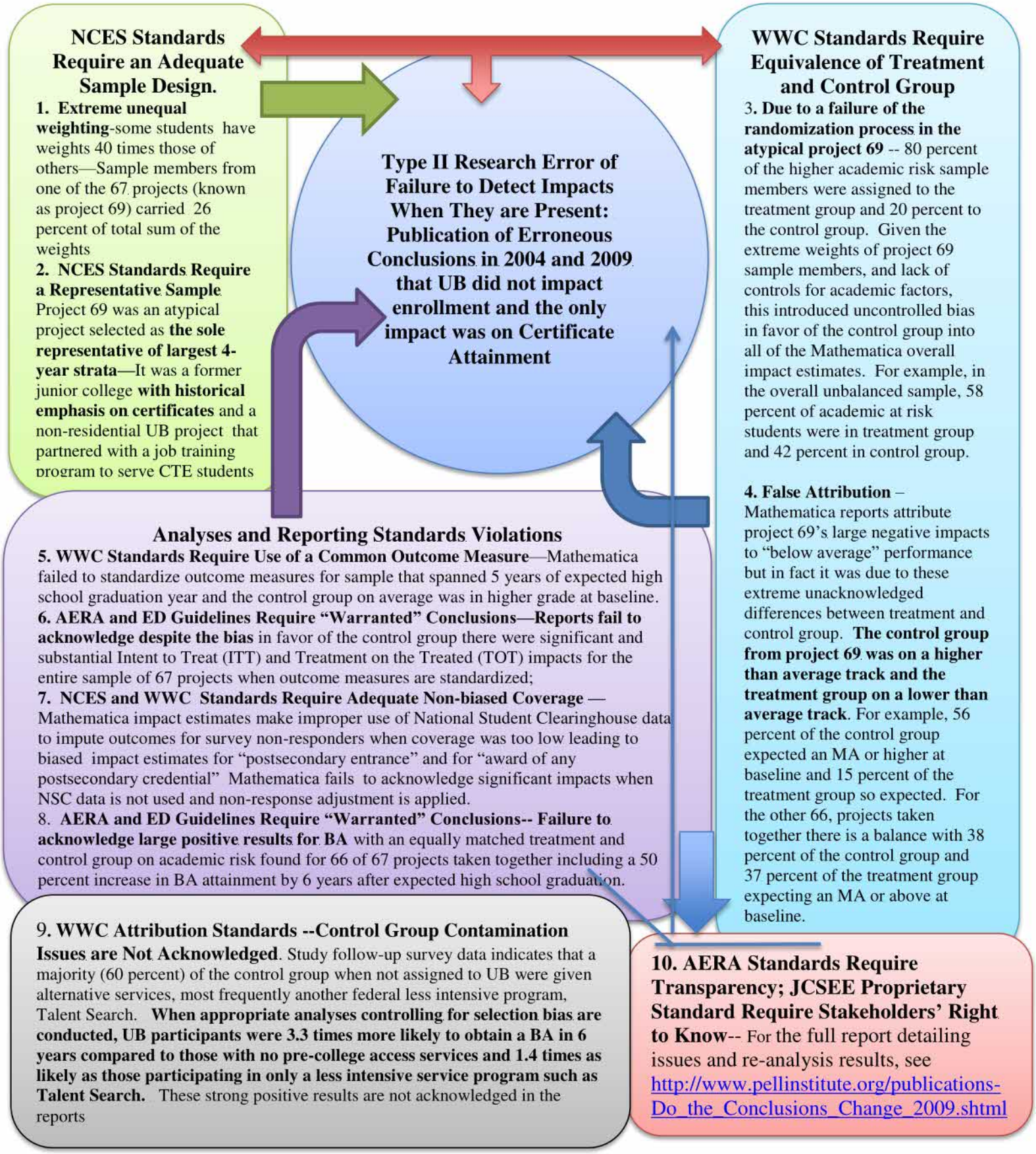
**Joint Committee on Standards for Education Evaluation Standards**: The Joint Committee Standards address ethics of research under the heading of Propriety. Standard P6 noted below discusses the full disclosure of findings

- **P6 Disclosure of Findings** The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results

**American Educational Research Association (AERA) Standards for Reporting on Empirical Social Science Research**

- Two overarching principles underlie the development of these reporting standards: the "sufficiency of the warrants" and the "transparency" of the report.

***Exhibit B-2. NCES, WWC, JCSEE, AERA Standards and ED Guidelines, Violations in the Mathematica Upward Bound (UB) Evaluation Reports***



**Violation 1**: **Using Flawed Sample Design to Make Inferences of Average Impact**. *NCES Statistical Standards and Department of Education Information Quality Guidelines require that the sampling follow correct procedures, that it be checked and found to be sufficient for robust estimation of the population for which the study is intended to generalize; (NCES Statistical Standards; commonly accepted procedures for inverse probability of selection weighting procedures).*

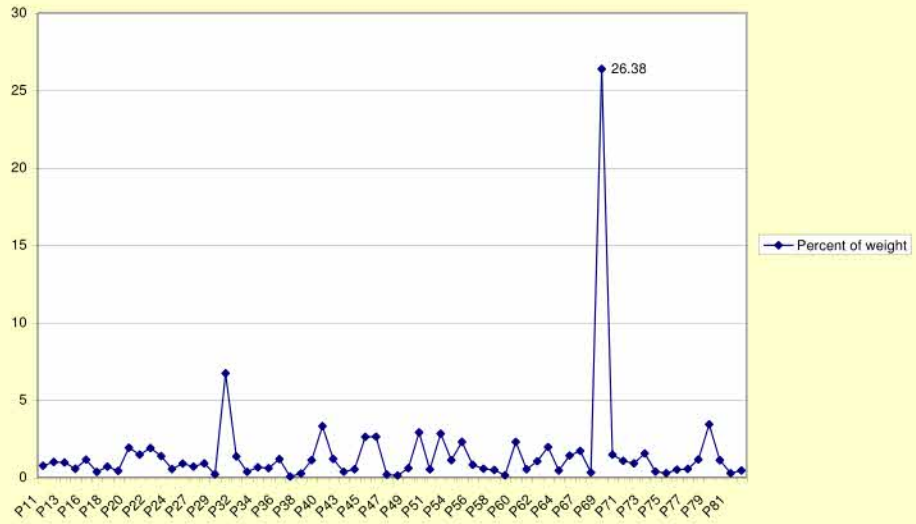
In a sample design that internal and external reviewers have consistently called "seriously flawed" and which violates NCES standards, the multi-stage sample design had only one single project selected to represent the largest study defined 4-year and above stratum. As Dr. Chromy has noted, "If representing the stratum of four-year institutions was a clearly stated objective of the study, than a sample of size 1 for this group is clearly inadequate." In the final stage of weighting this one project of 67 ended up carrying fully 26 percent of total student sample weights (Exhibit B-3). As noted, this meant that students sampled from this UB project had weights 40 times those of students in the lowest weighted UB projects.

An IES blind external reviewer summarized the issue with the following language asserting that inferential estimates could not be considered robust with such a design.

> *"The decision made in 1992-94 to select only a single project at random from this 26% share of the applicant population created a design in which design-based estimation and inferences for the full population could not be robust for the true population values. Simply applying a population weight to an inadequate sample of one cluster from a 26% share stratum will not correct this.*
>
> *What to do? With respect to design-based inferences for all other strata, the baseline sample of programs should enable robust inferences for that share of the UB survey population not included in the Project 69 stratum.". (IES external statistical reviewer C July 2008)*

All of the estimates in the 2004 and 2009 Mathematica reports were based on non-robust impact estimates that included Project 69. Impact estimates representing 74 percent of the intended population that had been identified as potentially more robust were not considered by Mathematica in the determination of conclusions they published. During preparation of the fifth and final UB report, PPSS Technical Monitors repeatedly recommended that Mathematica report UB impacts with and without project 69 and with and without the study weights and that the error issues with project 69 be considered in drawing conclusions about Upward Bound. These recommendations were repeatedly refused by Mathematica. The 2004 report, written before ED-PPSS knew of or raised questions concerning the limitations of the sample design, does not even mention the unequal weighting issue at all. The 2009 report does so only in a non-transparent manner that does not acknowledge the seriousness of the issue, or the equally significant if not more significant representational and lack of treatment and control group balance issues associated with project 69 discussed below in Violations 2 and 3.

**Exhibit B-3. Percentage distribution of sum of the weights by project for the 67 projects making up the study sample: National Evaluation of Upward Bound, study conducted 1992-93-2003-04**




**SOURCE:** Data tabulated April 2009 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education; study conducted 1992-93 to 2003-04.

**Violation 2. NCES Standards and ED Information Quality Guidelines Require a Representative Sample.** *The Mathematica reports use an "atypical project" as sole representative of largest public 4-year and above stratum. NCES standards specify that the sample be checked to see that it is representative of the population of interest.*

As PPSS career staff found out after the contract had ended, Mathematica had randomly selected "an atypical for its stratum" Upward Bound project to be the sole representative of the largest public stratum hosted at 4-year BA and above granting institutions. Dr. Cahalan and Dr. Goodwin together found when Mathematica finally sent the sampling frame list of eligible grantees to ED and the identity of project 69 became known to them at the end of the contract in December of 2007, that the project with 26 percent of the weight that was supposedly the sole representative of the largest 4-year stratum, although officially classified as a 4-year college was in fact a former minority serving junior college that had been taken over by a city college system to serve as their downtown campus. The UB program partnered with a job-training program and the grantee institution had historically had a large number of certificate programs linked to the CTE high schools served. It also did not have the hallmark 4-year grantees' UB summer residential program present in virtually all projects within the 4-year stratum it was representing—as it has no on-campus housing. The large diverse frame stratum for which project 69 is the sole representative included major flagship research universities that had UB grants at the time as well as public small 4-year majority white and majority black liberal arts colleges. None of these types of UB programs could be adequately represented by project 69.

This representation issue combined with the large weight and the serious lack of balance between the treatment and control group (to be discussed below) resulted in a probable over-estimate of the UB impact on certificate attainment and importantly a serious underestimate of the impact on BA receipt. In the context of recommending what should have been done at the start of the study external reviewer Dr. Chromy states:

> "With an imperfect sampling frame, it would be an accepted practice to check each project drawn and drop it from the sample if it does not meet the study population definition; this is a form of screening for eligibility. (Dr. James Chromy comments on the COE Request for Correction, October 2011)."

After the contract had ended, but while the final report was still under review, both Dr. Cahalan and Dr. Goodwin in separate communications with Mathematica asked Mathematia to acknowledge these concerns in the revisions to the report. Mathematica refused to do this and in fact sentences were inserted in the Executive Summary in the final changes negotiated with the political appointees that misled readers into thinking that project 69 was an adequate representative of the 4-year and above stratum. In the 2009 report Executive Summary Mathematica specifically states:

> *"Project 69 was similar to other projects in this stratum on a broad range of characteristics. Similarly data from the student surveys and NSC and FSA records indicated that the students from project 69 did not have unusual characteristics" (Executive Summary Mathematica Fifth Follow-up Report, Page xvii- xviii).*

The Mathematica 2009 report goes on to state that analyses that omit project 69:

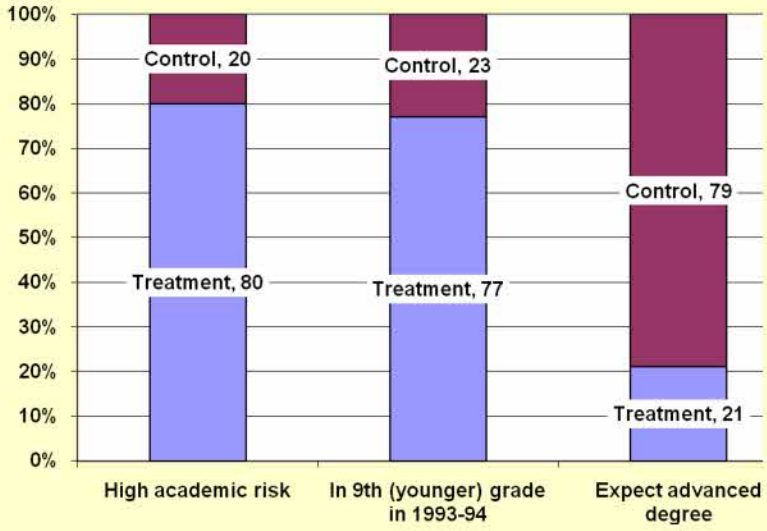
> *Do not appropriately represent the most common stratum of Upward Bound projects. Thus such analyses do not answer the evaluation's research questions about the impacts of the national Upward Bound program. Moreover the estimates for such analyses do not generalize to urban projects, large projects or any other well-defined subset of projects for which the findings might have policy implications. In contrast the findings from the main impact analyses, which include all projects weighted based on their selection probabilities are intended to generalize to the national Upward Bound program" (Executive Summary, Mathematica Fifth Follow-up Report, page xviii)*

**Violation 3: WWC Standards Require Equivalence of Treatment and Control Group**. *The Mathematica reports violate basic WWC and common random assignment study standards by using a seriously unbalanced non-equivalent treatment and control group to estimate Upward Bound impact and by not acknowledging the importance of this lack of balance in the report.*

WWC Standards require that the treatment and control group are equivalent on factors related to outcomes at the baseline and throughout the study. In what some external reviewers identified as a

probable failure to implement the random assignment correctly, PPSS found that the atypical and highly weighted project 69 also had a severe lack of balance between the treatment and control group. This lack of balance combined with the extreme weights resulted in an uncontrolled bias in favor of the control group in all of the Mathematica impact estimates as well as a lack of stability of estimates subject to non-response and non-coverage issues. As noted in project 69, the control group on average was found to be much more academically proficient, to be in a higher grade at baseline, and had much higher educational expectations than the treatment group (Exhibit B-4). Mathematica had no controls for academic risk indicators in their models and inadequate controls for grade at a fixed time and educational expectations. In addition as discussed under Violation 5, the utility of these later controls was compromised by the fact that outcome measures used were not standardized to expected high school graduation date. Due to the fact that students can enter Upward Bound from rising 9th graders to rising 12th graders, and the fact that recruitment for the study spanned 2 calendar year summers, the sample spanned 5 years of expected high school graduation year cohorts.

**Exhibit B-4. Percentage distributions for project 69 between treatment and control groups among those sample members who were a higher academic risk, in 9th (earlier) grade in 1993-94, and who expected an advanced degree at baseline: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (N= 85)**




Note large imbalance in project 69 distribution. Figure is read as follows: For example, among those who were classified as higher academic risk, 80 percent were in the treatment group and 20 percent in the control group. In a random assignment study distribution should be 50-50 between treatment and control group; figure shows imbalance in project 69.

**NOTE:** High academic risk includes those sample members in the bottom 20 percent of the sample on 9th grade GPA and other academic indictors. . There were not controls in the models for academic variables. **SOURCE:** Data tabulated April 2009 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education; study conducted 1992-93 to 2003-04.

The PPSS balance checks found that the treatment group sample resembled on average the students from an historically serving minority Career and Technical Education (CTE) Target high school served by the certificate granting project 69 institution grantee. On the other hand the control group from project 69 on average resembled the applicants to Upward Bound Math Science (UBMS) Initiative —known to be beginning at the time in the area under a separate UBMS grant (not selected to be in the study sample). PPSS suspected that the unusually large number of "applications" (baseline surveys submitted by this project) included baseline surveys for those more academically proficient and higher grade students interested in the newly initiated UBMS project. It is probable that these were not in fact "applicants" to the project 69 CTE focused Upward Bound program and should have been excluded from the study "waiting list sample" as ineligible.

Exhibit B-5 further illustrates the differences between the treatment and control group in project 69 with the treatment group resembling CTE students and the control group resembling on average students similar to Upward Bound Math Science applicants. For example, among the 66 other projects taken together about 38 percent of the control group and 37 percent of the treatment group reported at baseline that they expected an MA degree or higher. In contrast in project 69, among the control group 56 percent expected an MA degree or higher and among the treatment group 15 percent expected an MA degree or higher at baseline.

In project 69 the control group which consisted of all those who had completed a baseline survey and who were not randomly selected to be invited into the program as openings occurred was substantially larger than the treatment group. This was because the number of openings was less half the unusually large number of baseline surveys submitted by project 69. It should also be noted that within project 69 there were two project defined sub-strata so there was unequal weighting within project 69 itself, as most of the random assignment openings were filled from one of the two strata. This factor also contributed to the estimation issues from this project's sample. Whatever the reason they occurred, the substantial differences between the treatment and control group in project 69 accounted for the large negative impacts found in this project when it was considered alone.

**Exhibit B-5. Percentage of project 69 and all other projects having various attributes by treatment and control group status: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

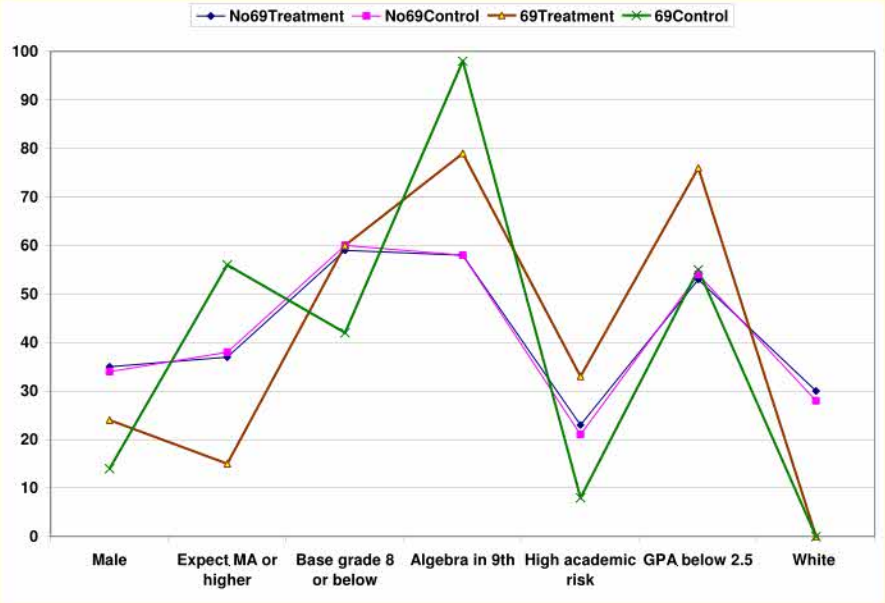


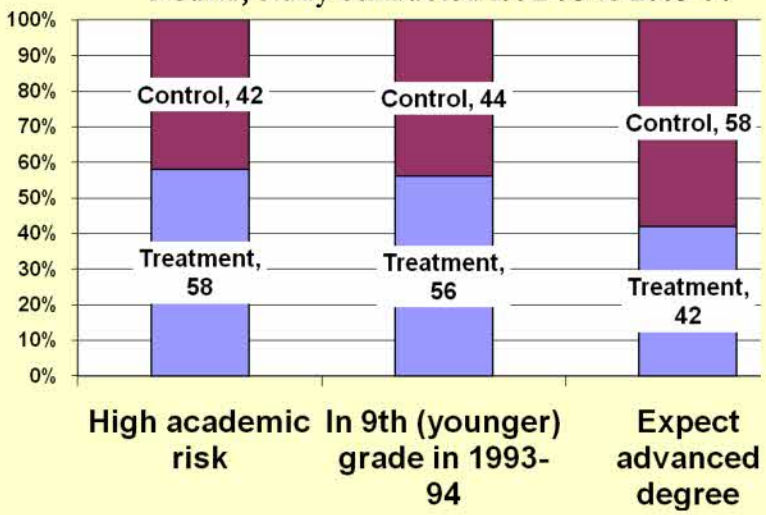

Figure shows that the UB treatment and control group are well matched without Project 69 on the variables in the chart; however, in project 69 the treatment and control group manifest substantial differences. For example, 56 percent of the control group in project 69 expected an MA or higher at baseline compared with 15 percent of the treatment group. In contrast, among the other 66 projects in the sample, 38 percent of the control group and 37 percent of the treatment group expected an MA or higher.

**NOTE:** Project 69 tabulation based on the 85 sample cases from project 69 (52 controls and 33 treatment cases -- poststratified weighted to 11,536 cases -- 5,768 treatment and 5,768 controls).. The category "No69treatment" and "No69control" represents all the other projects in the sample excluding project 69; these 66 projects are considered to represent 74 percent of the UB applicants in the study period.

**SOURCE:** Data tabulated December 2007 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education; study conducted 1992-93 to 2003-04.

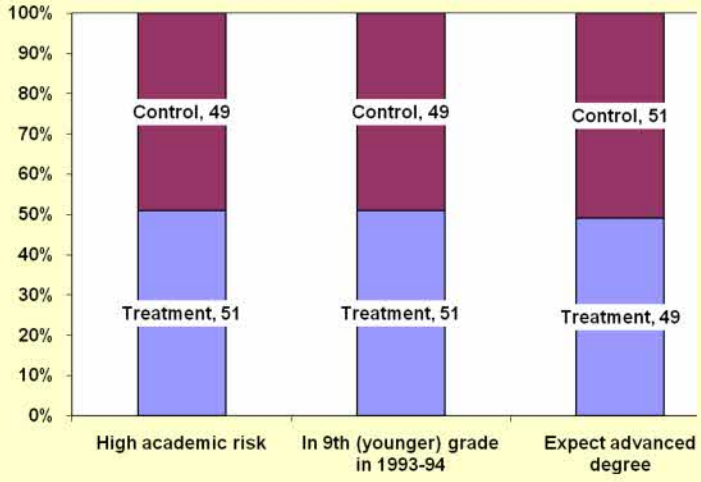
As noted, the academic risk differences between the treatment and control group were uncontrolled for in the analyses and introduced serious bias in favor of the control group into all of the Mathematica overall impact estimates (Exhibit B-6). As can be seen in Exhibit B-7, the other 66 projects when taken together exhibit a reasonable balance on these factors between the treatment and control group as one would expect in a random assignment study. For these reasons external reviewers such as Dr. Chromy and the IES blind external reviewer noted above recommended that the study acknowledge these limitations and attempt to make inferences only based on 66 of the 67 projects in the sample that exhibited a reasonable balance between treatment and control and a reasonable representation—especially in considering the impact on BA attainment.

**Exhibit B-6. Percentage distributions in all 67 sampled projects (including project 69) between treatment and control groups on various attributes: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**




Note with project 69 included **there is not a balance or equivalence between treatment and control group** as WWC standards require in a random assignment study. Figure is read as follows: For example, among those who were classified as higher academic risk, 58 percent were in the treatment group and 42 percent in the control group. In a random assignment study distribution should be about 50-50 between treatment and control group; figure shows imbalance in overall sample with project 69 included.

**Exhibit B-7. Percentage distributions for 66 of 67 sampled projects (excluding project 69) between treatment and control groups among those sample members who were a higher academic risk, in 9th (earlier) grade in 1993-94, and who expected an advanced degree at baseline: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**




Note without project 69 there is a balance between treatment and control group as expected in random assignment study. Figure is read as follows: For example, among those who were classified as higher academic risk, 51 percent were in the treatment group and 49 percent in the control group. In a random assignment study distribution should be about 50-50 between treatment and control group.

**NOTE:** High academic risk includes those sample members in the bottom 20 percent of the sample on 9th grade GPA and other academic indictors. . There were not controls in the models for academic variables

**SOURCE:** Data tabulated April 2009 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education; study conducted 1992-93 to 2003-04.

**Violation 4: False Attribution of Project 69's negative impacts when considered alone ---***What Works Clearinghouse Standards require that the impacts be attributable to the intervention being studied***.**

Mathematica's failure to detect impacts and type II errors are related to a number of factors, but as seen in Exhibits B-4 to B-7, one factor in their error was attributing a lack of impact findings as related to the Upward Bound program performance and not to the uncontrolled bias present due to the lack of balance between the treatment and control group on academic factors. The Mathematica reports specifically has misleading attribution of "below average impacts" for project 69, in implying that project 69 had below average impacts due to the Upward Bound program's "below average" (poor) performance when in fact these negative impacts were due to the gross lack of equivalency between the treatment and control group in this project. The Mathematica report states: "*Because Project 69 had below average impacts, reducing its weight relative to other projects resulted in larger overall impacts for most outcomes compared with the findings from the main impact analysis, which weighted all sample members according to their actual selection probabilities."* This is a misleading statement about the effectiveness of project 69. As noted above in Exhibit B-4, a closer look at project 69's treatment and control group clearly reveals that the so-called "below average" (negative impacts) in this project were not due to project 69's "poor performance" but were due in fact to the extreme differences between the treatment and control group in favor of the control group in this project.

As noted above, the facts concerning project 69's representational issues and lack of treatment and control group balance issues are not acknowledged in the Mathematica reports. This misrepresentation of a largely 2-year and less than 2-year grantee as the sole representative of the largest 4-year public stratum, combined with the extreme large weight and the uncontrolled for academic factor bias in favor of the control group in this project's sample contributes to a probable type I error of over-estimating the impact of UB on the attainment of certificates of the type awarded by the project 69 grantee institution and a type II error of failure to detect effects for the attainment of Bachelor's degrees. Very strong positive effects on BA attainment were found for 66 of the 67 sampled projects taken together, which were found to meet WWC standards for baseline equivalence and were well matched when taken together on relevant attributes (see discussion below on BA attainment), but not when project 69 is included in the impact estimates.

**Violation 5. Failure to Standardize Outcome Measures.** *WWC and NCES standards require that any comparisons use standardized and precise common outcome measure. Mathematica did not do this in any of its outcome measures used to assess UB impact.*

As noted, the UB student sample spanned 5 years of expected high school graduation years. When PPSS reviewers asked that Mathematica standardize the outcome measures, they argued that their random assignment method made standardizing by expected high school graduation year unnecessary. Consequently Mathematica failed to standardize their outcome measures to differences in expected high school graduation years in either the 2004 or 2009 reports. PPSS QA review found the control group on average to be in a higher grade at a fixed time period, and this contributed to the uncontrolled bias in the Mathematica impact estimates for key postsecondary outcome measures (enrollment, financial aid,

postsecondary degrees and credentials). The lack of standardization of outcome measures also decreased the utility of the control variables used in the regression models. After Mathematica did not respond positively to PPSS and OPE's suggestions that the outcome measures be standardized and after the contract had ended, in spring of 2008, PPSS internal staff derived the dates necessary to standardize the primary outcome measures. Exhibit B-8 below shows the differences between unstandardized impact estimates and standardized impact estimates using the third follow-up survey data and administrative data.

**Exhibit B-8. Various model results using third follow-up survey responders only and using full longitudinal sample for evidence of entering postsecondary for ITT and TOT models: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| | **All sampling strata** | | **One project with bias removed (the remainder represents 74 percent of Horizons waiting list)** | |
|---|---|---|---|---|
| | **Given Opportunity (ITT)** | **Participated in UB/UBMS (TOT/CACE)** | **Given Opportunity (ITT)** | **Participated in UB/UBMS (TOT/CACE)** |
| Third follow-up survey responders only with no administrative records and no standardization of outcome to expected high school graduation year (EHSGY); uses non-response adjusted weight | pr-T = 76.4<br>pr-C = 75.4<br>Difference = 1.0NS | xb T = 75.4<br>xb C = 71.7<br>Difference = 3.7 NS | pr T = 77.8<br>Pr C = 72.2<br>Difference = 5.7** | xb T = 77.6<br>xb C = 67.7<br>Difference = 9.9* |
| Third follow-up survey responders only with no administrative records or other applicable surveys, but with standardization to +1 (18 months) of expected high school graduation year; uses non-response adjusted weight | pr-T = 71.2<br>pr-C = 68.2<br>Difference = 3.0 NS | xb T = 71.4<br>xb C = 65.2<br>Difference = 6.1 NS | pr T = 73.3<br>Pr C = 65.8<br>Difference = 7.5*** | xb T = 74.0<br>xb C = 61.9<br>Difference = 12.1*** |
| Third follow-up survey responders only – standardized to +1 (18months) of EHSGY and uses all applicable surveys and Student Financial Aid (SFA) records; uses non-response adjusted weight | pr-T = 75.9<br>pr-C = 71.4<br>Difference = 4.6* | xb T = 76.0<br>xb C = 67.8<br>Difference = 8.2 NS.11 | pr T = 77.8<br>Pr C = 70.0<br>Difference = 7.8**** | xb T = 78.2<br>xb C = 65.6<br>Difference = 12.6*** |
| Includes all sample members, standardized to +1 (18months) of EHSGY and uses all applicable surveys and SFA records; uses poststratified adjusted weight | pr-T = 72.9<br>pr-C = 66.0<br>Difference = 6.9**** | xb T = 73.5<br>xb C = 62.5<br>Difference = 10.9**** | pr T = 73.3<br>Pr C = 64.3<br>Difference = 9.1*** | xb T = 74.6<br>xb C = 60.4<br>Difference = 14.2**** |

*/**/***/**** Significant at 0.10/0.05/.01/00 level; NS = not significant at the .10 level or below. UB = regular Upward Bound; UBMS = Upward Bound Math/Science; ITT = intent to treat; TOT = treated on treated; CACE = complier average causal effect; T = treatment; C = control or comparison; pr = estimated probability from STATA logit regression; xb = linear prediction from STATA ivreg (instrumental variables regression). **SOURCE:** Data tabulated December 2007 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), US Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

The highly influential Third Follow-Up report (Myers, et. al 2004) reporting no impacts on postsecondary outcomes for Upward Bound did not standardized outcomes by expected high school graduation year or use administrative records in addition to the surveys. All impact estimates included the biases introduced by project 69 and no estimates without project 69 were reported. Nor did the report acknowledge that the "no impact" estimates were being driven by one project with large negative impacts. The 2004 report did not reveal any of the bias issues surrounding project 69's extreme unequal weighting, representational issues or the non-equivalent treatment and control group resulting in uncontrolled bias in favor of the control group. (the issues discussed above in violations 1 to 4). The Third Follow-up findings or lack of findings formed the formal justification of the OMB PART rating of "Ineffective" and the Bush Administration's budget requests in FY2005 and FY2006 to zero fund Upward Bound, Upward Bound Math Science, Talent Search and GEAR UP.[5]

**Violation 6: Failure to Acknowledge and Report Positive Findings When Standardization Was Implemented.** *AERA Standards, ED Information Quality Guidelines, and Joint Committee Standards for Education Evaluation (JCSEE) all require that conclusions be "warranted"*

PPSS review comments to the draft fifth follow-up report in 2007 (Sefter et.al. 2009) had asked for standardization of the outcome measures to expected high school graduation cohort year. However Mathematica continued to argue that this was unnecessary due to their randomization process. As noted in response in early 2008, after the contract ended PPSS internal staff did the standardization and sent a paper to Mathematica in the spring detailing the results that included the output from the STATA runs. Examples of these positive impact estimates are graphed in Exhibits B-9 and B-10 and documentation of the output from the regression models is included Attachment C. Mathematica choose to ignore these positive findings and to engage in a debate about the Baseline Survey variable PPSS used to derive expected high school graduation year. In response PPSS used a different survey variable from the First Follow up Survey and found the same significant and substantial positive impact estimate results which are also included in Attachment C.[6]

The PPSS impact estimates using two different methods to estimate expected high school graduation year both showed significant and substantive positive impacts for Upward Bound with and without project 69 for enrollment and application for and award of financial aid . As noted, despite this information Mathematica chose to not report or even acknowledge that these documented findings had been sent to them in their revisions to the Fifth Follow-up report. Moreover, despite the fact that both Dr. Cahalan and Dr. Goodwin had repeatedly questioned Mathematica's methods and conclusions, Mathematica thanked Dr. Goodwin and Dr. Cahalan in the acknowledgements to the report with no mention of the fact that both Dr. Goodwin and Dr. Cahalan had communicated to Mathematica that they

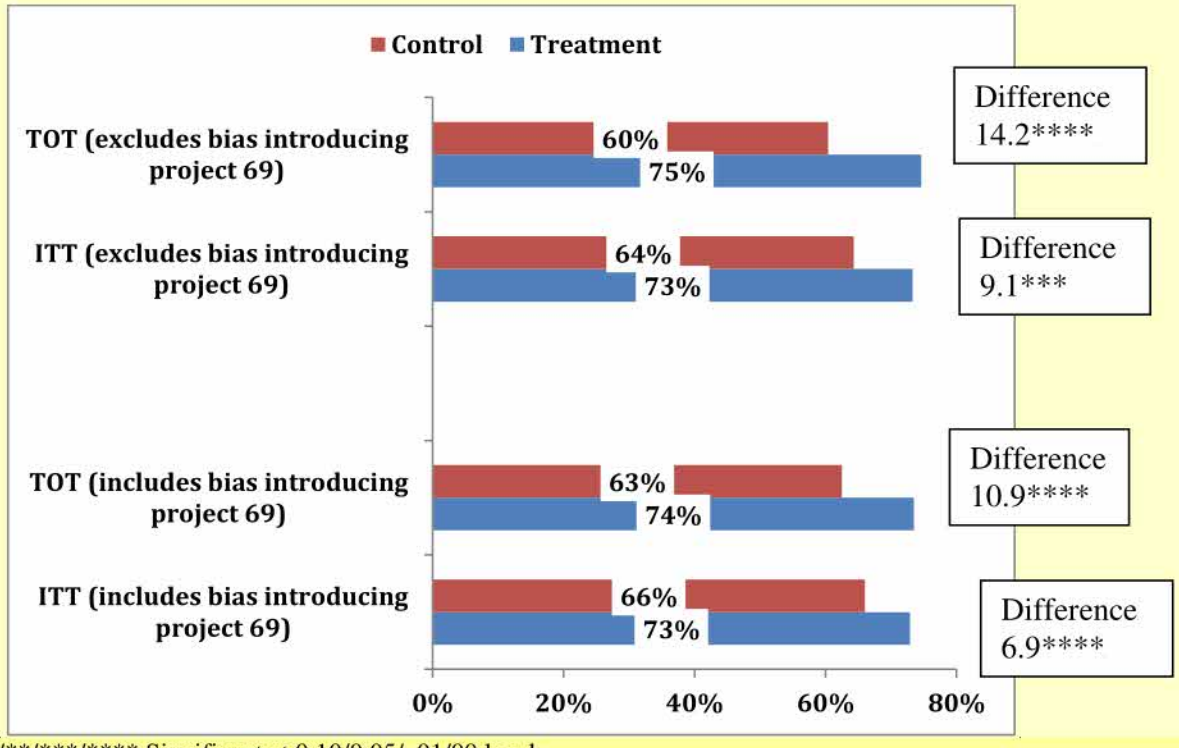
[5] The Third Follow Up results of no overall impact but very large impacts on postsecondary entrance for those students classified as the highest academic risk, also led to the efforts of OMB and the Department of Education to re-focus the Upward Bound program away from preparation for BA attainment and toward serving those students who were considered to be the most at academic risk in their high schools (defined as students who had failed a high school competency test or had very low GPAs) though the Upward Bound Initiative.

[6] Attachment C is taken from appendices from a paper prepared by Cahalan that includes impact estimates from using two different variables to ascertain expected high school graduation year. Cahalan found that the impact estimates were similar when the two different survey questions were used. This exercise was undertaken when Mathematica questioned the variable used by PPSS to ascertain expected graduation year.

viewed the draft report and its conclusions as "seriously flawed." This lack of mention of the positive impacts when results were standardized and lack of mention of the PPSS monitoring staff concerns, gave the misleading impression to readers and reviewers that the concerns with the report conclusions raised by PPSS were unfounded or had been adequately addressed. As can be seen from the examples in Exhibits B-8 to B-10 there is clear documented evidence of positive postsecondary impact when results are standardized by expected high school graduation year for ITT and TOT estimates with and without the bias introducing project 69.

Exhibit B-9 summarizes standardized results for enrollment within +1 year after expected high school graduation year. Similar findings were found for enrollment in +4 year after high school. Exhibit B-10 presents federal aid application within 4 years of expected high school graduation year and is based solely on the administrative records from 10 years of federal aid files (not subject to survey or coverage response bias). Contrary to what is reported in the Mathematica fifth follow up report there is clear evidence of impact on both application and award of aid when aid results from 10 years of federal aid files are standardized by expected high school graduation year.

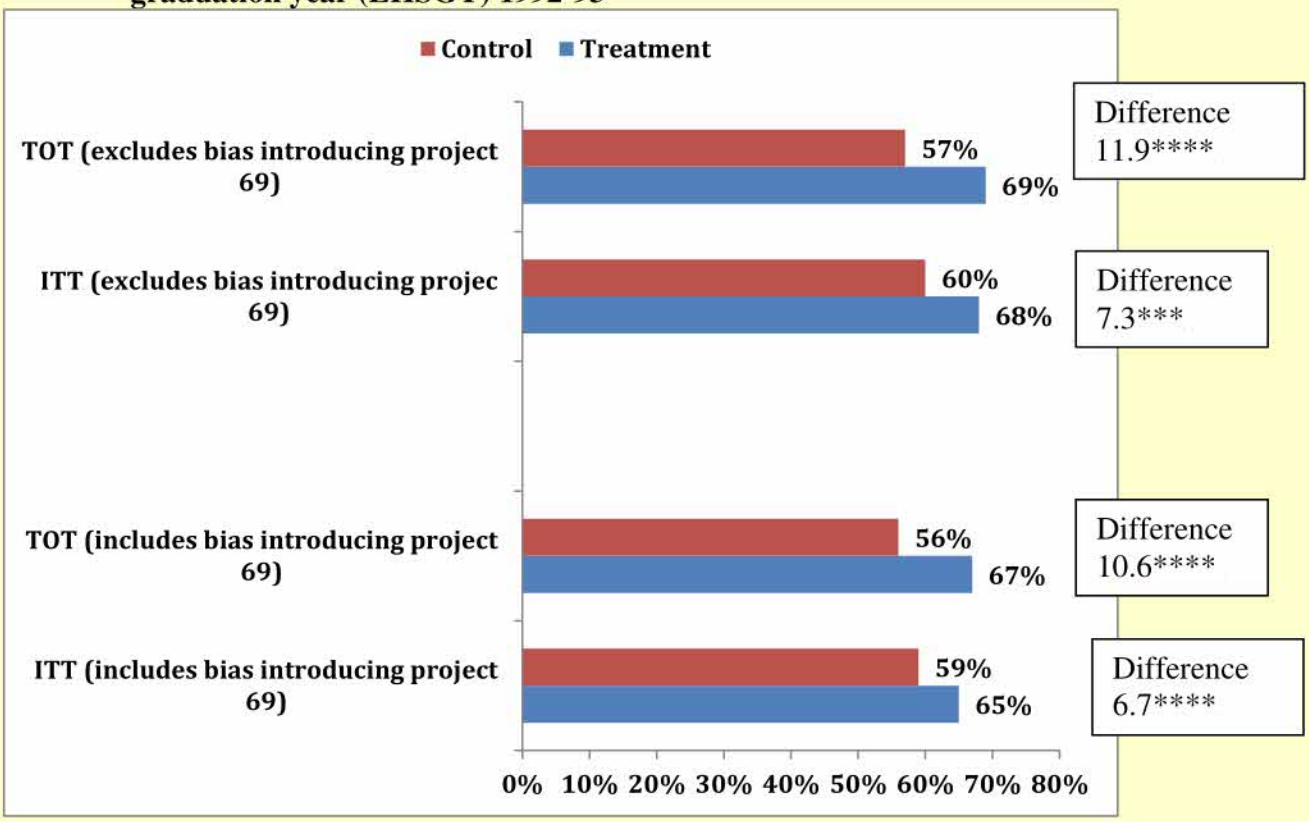
**Exhibit B-9. Treated on the Treated (TOT) and Intent to Treat (ITT) estimates of impact of Upward Bound (UB) on postsecondary entrance within +1 year (18 months) of expected high school graduation year (EHSGY) 1992-93 to 2003-04**




*/**/***/**** Significant at 0.10/0.05/. 01/00 level.

**NOTE**. Model based estimates based on STATA logistic and instrumental variables regression and also taking into account the complex sample design. Based on responses to 5 follow-up surveys and federal student aid files. **SOURCE:** Data tabulated January 2008 using:. National Evaluation of Upward Bound data files, and federal Student Financial Aid (SFA) files 1994-95 to 2003-04. (Excerpted from the *Cahalan Re-Analysis Report*, Figure IV).

**Exhibit B-10. Treated on the Treated (TOT) and Intent to Treat (ITT) estimates of impact of Upward Bound (UB) on federal aid application within +4 years of expected high school graduation year (EHSGY) 1992-93**




*/**/***/**** Significant at 0.10/0.05/. 01/00 level;;**NOTE:** Estimated rates from STATA logistic and instrumental variables regression taking into account the complex sample design. Weighted data use poststratified weights. Based on 10 years of financial aid administrative record data. . . . **SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04

**Violation 7. Coverage and Bias Issues in the Improper uses of National Student Clearinghouse (NSC) data.** *Both NCES and WWC standards affirm that the study must achieve adequate overall and differential coverage and response, and must concern itself with overall attrition and with differential attrition, non-response and non-coverage bias issues for all of the data sources used relative to the population of interest.*

In the fifth follow-up report Mathematica violated NCES coverage standards in making improper use of the NSC data to impute all fifth follow-up survey non-responders (about 25 percent of the sample) not found on the NSC files to negative values for the outcomes of postsecondary enrollment, and award of any degrees. The applicable time period was when NSC reported only 26 percent coverage for enrollment and when 2-year and below degree data was not yet even being collected by NSC. There was also evidence of bias due to the fact that the heavily weighted project 69 did not participate in NSC

until a date after its sampled target school graduates would have graduated high school. There is evidence that this improper use of the NSC data for the fifth follow-up non-responders combined with the unequal weights increased sensitivity to small variations in coverage and response and led to erroneous conclusions from the study for enrollment estimates but also especially for the conclusions with regard to the important output measure of "award of any postsecondary degree or certificate."

As the statement below taken from the Executive Summary notes, Mathematica reported that the study detected no impact on award of "any postsecondary degree or credential." The 2009 Mathematica report executive summary states:

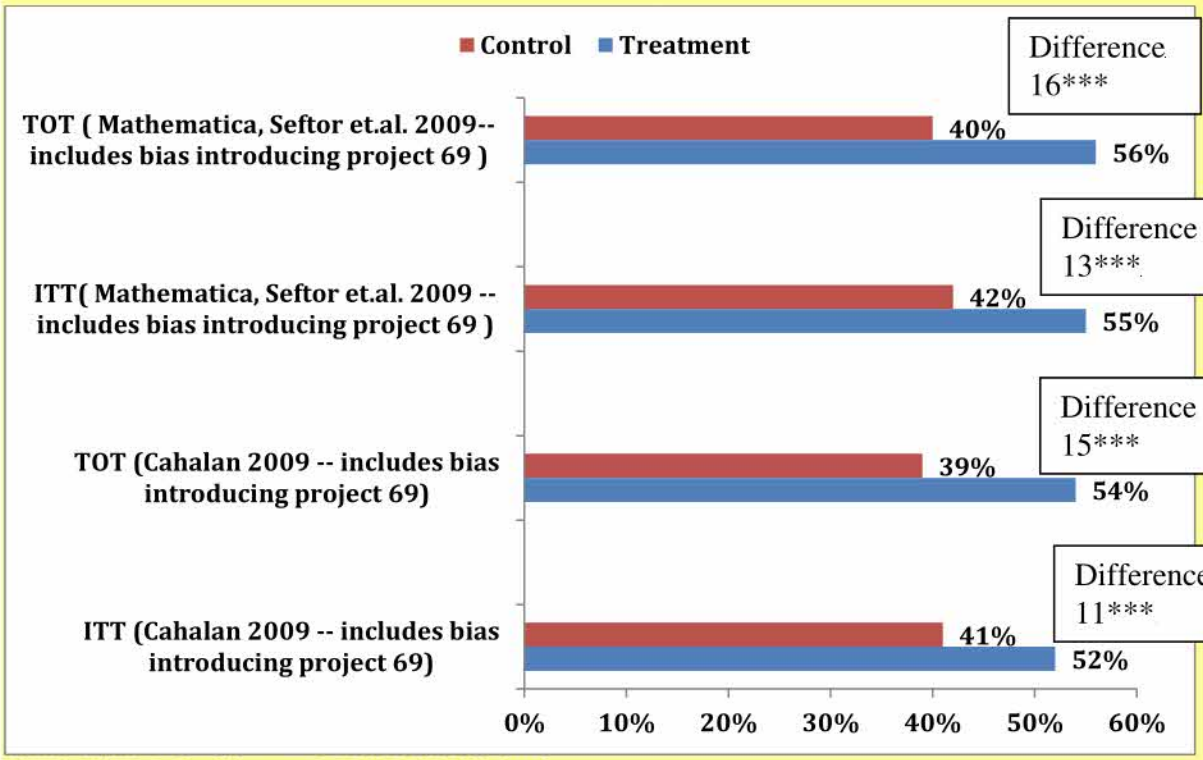
> ***The impacts on receiving any postsecondary credential and receiving a bachelor's degree are 2 and 0 percentage points (effect size = 5 and 0 percent), respectively, and are not statistically significant.***

In making this statement, Mathematica chose to ignore the statistically significant and substantial impacts the study detected on "award of any postsecondary degree or credential" by the end of the study period based on responses to the fifth follow up survey adjusted for non-response without use of the NSC data. These impacts are included in the Mathematica final report in Appendix Tables C7 and C14 (Seftor et.al 2009). These substantial positive findings went unacknowledged in the Mathematica conclusions reported in the body of the report concerning Upward Bound..[7] This data is graphed in Exhibit B-11. As can be seen, Mathematica's own estimate of attainment of "any postsecondary degree or credential" based on responders to the Fifth-Follow-Up Survey shows a positive substantial and significant Intent To Treat (ITT) impact of UB on award of "Any postsecondary degree or credential" of 13 percentage points (55 percent for UB and 42 percent for the control group) and a Treatment On the Treated (TOT) estimate of a 16 percentage point difference. Similar positive UB impacts on award of any postsecondary degree or certificate were reported by Cahalan 2009 available at (see http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml ) and are also included in Exhibit B-11. **As noted these large and statistically significant positive impacts, tabulated in their own analyses, were ignored by Mathematica in their text discussion of impact and went unacknowledged in their widely quoted conclusions concerning the "efficacy of Upward Bound". As can be seen above, the Mathematica Executive Summary specifically falsely reports that the study did not detect statistically significant impacts on receiving "any postsecondary credential."** Had Mathematica followed the same procedures as they followed in the 2004 (Third Follow-up) report and the unpublished draft fourth follow-up report basing conclusions on the follow-up survey data adjusted for non-response these would have been the results upon which they based conclusions. However, against PPSS repeated advice to only use NSC data cautiously for BA degree which would have occurred in a later period, as well as that of external IES reviewers to be conservative in use of NSC data Mathematica chose to only put into text tables information with the NSC data that coded non-responders to the fifth follow-up survey without degree information on the NSC as not

[7] Due to positive survey bias, it is possible that the estimates based on survey responders even adjusted for non-response including project 69 overestimates total degree-certificate attainment by the sample, however there were not differences in response rates between the treatment and control group for this round of the surveys—about 75 percent in the 5th follow-up.

having any degree. This violation of NCES coverage standards combined with the bias and instability from project 69's weights and the lack of treatment control group equivalency on academic measures led to a Type II error of failure to report impacts for Upward Bound on award of "any postsecondary degree or credential."

**Exhibit B-11: Treatment on the Treated (TOT) and Intent to Treat (ITT) and impact estimates for outcome measure of Award of Any Postsecondary Degree or Certificate by the end of the study period based on 67 of 67 sampled projects respondents to the Fifth Follow-Up Survey**




*/**/***/**** Significant at 0.10/0.05/.01/00 level.

NOTE: **Based on 67 of 67 projects sampled**. TOT = Treatment on the Treated; ITT= Intent to Treat **NOTE:** Estimated rates from STATA logistic and instrumental variables regression taking into account the complex sample design. Weighted data use poststratified weights.. Cahalan impact estimates used a non-response adjusted weight prepared by Mathematica. Mathematica impacts taken from Appendix Table C-7 and C-14 in the Seftor et.al. 2009 report and are not acknowledged in conclusions reported by Mathematica.

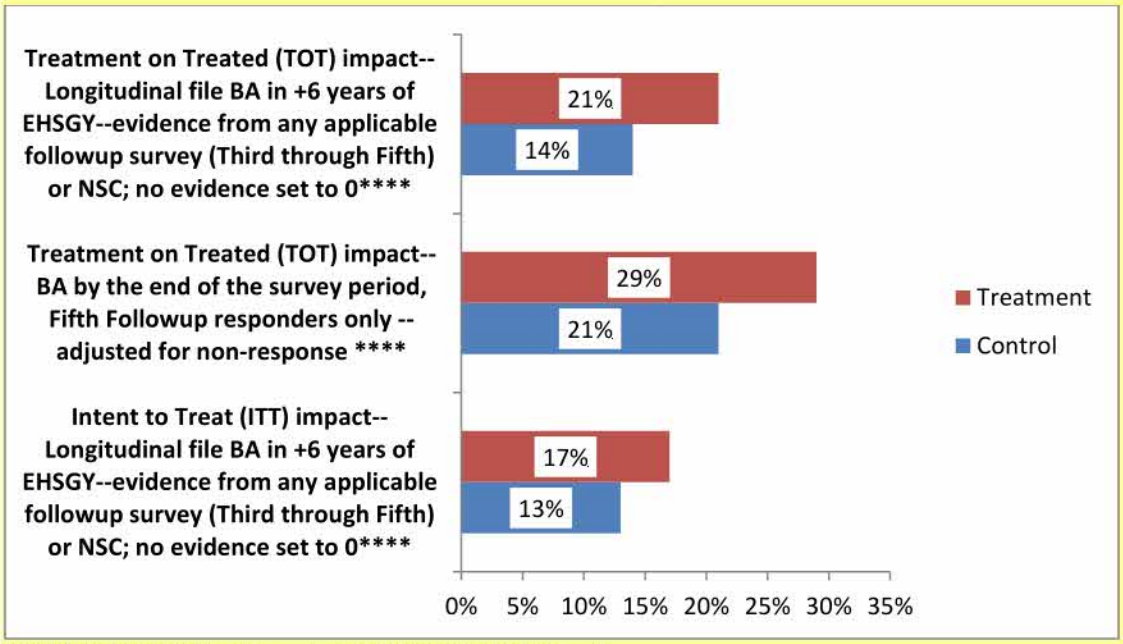
**SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04

**Violation 8: Failure to Report and Acknowledge Large Positive Impacts on BA Attainment for the 66 of the 67 sampled projects that when taken together had an equally balanced treatment and control group on academic risk factors.**

As noted above, Mathematica reported no impact for Upward Bound on BA attainment. Preparation for and fostering BA attainment has historically been a major UB programmatic focus. The UB summer and academic year programs by legislative mandate have a strong academic focus and encourage 4-year

college enrollment whenever possible.. Given the representational issues of the heavily weighted project 69 (discussed under violations 1 and 2) combined with the fact that at baseline the treatment group from project 69 was on average on track for certificates or two-year college credentials, and the control group from project 69 was on average on track for advanced degree attainment, (discussed under violations 3 and 4) no impacts were found with or without standardization when the bias introducing project 69 was included in the overall impact estimates. However, PPSS found very strong impacts for the 66 of 67 projects that had a balanced treatment and control group when taken together. None of these more robust estimates for a balanced treatment and control group are acknowledged in the Mathematica reports. Exhibit B-12 summarizes impacts found and reported by the PPSS QA review and sent to Mathematica during the review process for their report. These estimates were ignored in the Mathematica conclusions and reporting that UB had no impact on BA attainment.. For the Treatment on the Treated (TOT) impact PPSS found that UB participation increased BA attainment by 50 percent (from 14 to 21 percent), and for the ITT estimate BA attainment was increased by 28 percent.

**Exhibit B-12. Impact of Upward Bound (UB) on Bachelor's (BA) degree attainment: estimates based on 66 of 67 projects in UB sample: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**




*/**/***/**** Significant at 0.10/0.05/.01/00 level.

NOTE: TOT = Treatment on the Treated; ITT= Intent to Treat; EHSGY = Expected High School Graduation Year; NSC = National Student Clearinghouse; Estimates based on 66 of 67 projects in sample representing 74 percent of UB at the time of the study. One project removed due to introducing bias into estimates in favor of the control group and representational issues. Model based estimates based on STATA logistic and instrumental variables regression taking into account the complex sample design. We use a 2-stage instrumental variables regression procedure to control for selection effects for the Treatment on the Treated (TOT) impact estimates. ITT estimates include 14 percent of control group who were in Upward Bound Math Science or UB and 20-26 percent of treatment group who did not enter Upward Bound. Calculated January 2010.

**Violation 9: Attribution Error. Failure to Adequately Acknowledge Issues with Control Group Service Substitution and Treatment Group Waiting List Dropping Out.** *What Works Clearinghouse Standards require that the intervention whose effects are being measured can be clearly attributed to the intervention and that the only difference between the treatment and control group is the intervention. Generally accepted research standards require that the treatment and control group are treated equally except for the treatment; and the treatment and control group are mutually exclusive with regard to the treatment.*

One of the most difficult challenges of education random assignment studies, especially of voluntary support service federal programs, concerns establishing and maintaining clearly distinguished treatment and control groups. This issue has been repeatedly raised by UB study project participants and stakeholders concerning the Mathematica Upward Bound evaluation from the period of the initial random assignment process. This issue also formed the basis of the arguments made in Congress against a new UB evaluation study begun in late 2006, and cancelled by ED in early 2008 following Congressional cutting off of funding. It was argued that it would be unethical to purposively increase recruitment among the 9th graders targeted in the study (those who were most academically deficient and most vulnerable), and then to limit entrance into UB program throughout high school to half of those recruited. If time sensitive services were not denied and alternative services were provided then the study results might be confounded by control group substitution and treatment group dropping out.

This type of confounding was the situation of the Mathematica Upward Bound study. The design was not a randomized control trial in the classic sense—in that the study did not attempt to control, and could not have ethically controlled, student's participating in time sensitive similar pre-college supplemental services from other sources. In the specific two summer window of years, only the opportunity to be invited to apply to the specific regular UB program sampled was done in a randomized manner from a group of middle and early high school students who completed a baseline survey indicating interest in the program. TRIO pre-college programs (Upward Bound, Upward Bound Math Science and Talent Search) in a given area often work together to serve the same target schools with large percentages of students meeting the legislatively mandated eligibility requirements---low-income, disabled, and first generation students. As the data below indicates, a large majority of those students in the control group not selected to be given the regular Upward Bound opportunity, were then given the opportunity for Talent Search, or in some cases Upward Bound Math Science, a new initiative being organized on a regional basis in this period.

The Mathematica Upward Bound study baseline and follow-up surveys contained questions (sometimes quite detailed) about other pre-college support or supplemental service participation, although these questions were somewhat different in each of the applicable survey rounds and are limited by the fact that the students were also in different grades at the time they completed the various survey rounds. They also suffer from the fact that the control group was not asked directly about any regular UB participation.[8] However, sufficient information was collected to classify whether the student reported any other pre-college support or supplemental services, and whether the study participant participated

[8] In asking about alternative services, control group members were given a list of specific and general programs. For the control group the list did not contain the name "UB program." They were asked about UBMS participation.

in UBMS.[9] This information, summarized in Exhibit B-13, can be used to gain some understanding of how much of an issue equivalent and/or similar service receipt was for this study.

**Exhibit B-13. Number and percent of study sample participating in UB or UBMS and other pre-college support or supplemental service programs with academic components, by treatment and control group status: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| | Random Assigned Treatment | | Random Assigned Control | | Total Horizons Study | |
|---|---|---|---|---|---|---|
| | **Unweighted** | **Poststratified Weighted** | **Unweighted** | **Poststratified Weighted** | **Unweighted** | **Poststratified Weighted** |
| **Total** | 1,524 (100%) | 21,866 (100%) | 1,320 (100%) | 21,866 (100%) | 2,844 (100%) | 43,732 (100%) |
| **Reported participated in UB or UBMS service** | 1,247 (82%) | 17,843 (82%) | 180 (14%) | 2702 (12%) | 1,427 (50%) | 20,545 (47%) |
| **Reported participated in "another" (not UB and not UBMS) pre-college support or supplemental service program only** | 128 (8%) | 2,332 (11%) | 618 (47%) | 10,513 (48%) | 746(26%) | 12,845 (29%) |
| **Did not report participation in any type of (UB, UBMS, or other) pre-college support or supplemental service program** | 149 (10%) | 1690 (8%) | 522 (40%) | 8651 (40%) | 671 (24%) | 10,342 (24%) |
| **Reported participated in any type (UB, UBMS, or other) of pre-college support or supplemental service program** | 1375 (90%) | 20,176 (92%) | 798 (61%) | 13,215 (60%) | 2173 (76%) | 33,390 (76%) |

**NOTE:** Percentages given in parentheses. UB = Upward Bound; UBMS = Upward Bound Math/Science. Weighted data use poststratified weights for longitudinal file. **SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

[9] Information was collected on the surveys about length of participation and type of programs on the various surveys that could be analyzed in more detail.

**Participation in the UB Program or Upward Bound Math Science by the Treatment and Control Group**. About 26 percent of the Treatment Group maintained in the ITT analysis which formed the basis of Mathematica's conclusions was coded as "waiting list" dropouts during the period in which cases were randomly selected to be recruited to apply for entrance into the program, and about 20 percent of the treatment sample reported on the First Follow-up that they did not enter regular Upward Bound.[10] A portion of this 20 percent reported they could not remember being given the opportunity when asked about it a year later. Conversely about 12 to 14 percent of the control group reported they entered into Upward Bound Math Science (UBMS) or Upward Bound. The Mathematica Fifth-Follow-up Report, while emphasizing Intent to Treat (ITT), includes some Treatment on the Treated (TOT) analysis taking into account the 12-14 percent UBMS crossovers and the treatment non-UB participants (unlike the Third Follow-up report which did not recognize UBMS participation by the control group as a crossover). However as one of the IES external reviewers noted given that one fourth to one-fifth of the so called "treatment group" did not enter UB and 12-14 percent of the control group was in UBMS or UB, the TOT estimates may be more meaningful statistic for this study. In the study year procedures were altered to ensure there would be double the number of baseline survey completers as openings. As noted, instead of obtaining actual applications to UB, those who completed the baseline surveys were considered on a "waiting list" for participation and in the study years no one could get on the UB "waiting list" without completing the baseline survey. All of the students were minors and over half were in middle school when completing the baseline survey; hence their actual entry into the UB program that next summer which was typically a residential program was related to parental permissions and family mobility. Low income families have high levels of mobility.

**Participation in other lessor pre-college services**. Examination of study survey data also revealed that a majority (60 percent) of the control group reported participating in some form of supplemental pre-college programs by the end of high school (Exhibit B-13). Frequently when students were not randomly selected for Upward Bound, they were placed in Talent Search or another similar pre-college program. Presumably most of these programs were less intensive than Upward Bound. Cahalan reports that PPSS requested that Mathematica use the information from the baseline and follow-up surveys on alternative service receipt to statistically address issues of service substitution and treatment group non-entry into UB; however, Mathematica declined to conduct these analyses.

This is the same issue addressed by noble laureate James Heckman, and co-authors (Heckman Hohman, Smith, and Khoo 2000) re-analysis of the Job Training Partnership Act (JTPA) evaluation in which they considered the interpretation of evidence from social experiments when persons randomized out of a program being evaluated have good substitutes for it, and when persons randomized into a program do

[10] There was also a group of about 192 "waiting list participants" who completed baseline surveys who were excluded from the analyses and their weights re-distributed among those randomly assigned because they were selected with certainty into the group being given the UB opportunity. Projects were allowed to serve selected students in Upward Bound in the study period to whom they might have made a prior commitment or if the projects wished to serve them for group cohesion or diversity purposes.

not enter the program or drop out. Using data from an experimental evaluation of JTPA classroom-training programs, they documented the empirical importance of control group substitution and treatment group dropping out. They note that "evidence that one program is ineffective relative to close substitutes is not evidence that the type of service provided by all of the programs is ineffective, although that is the way experimental evidence is often interpreted" (Heckman et. al. 2000).

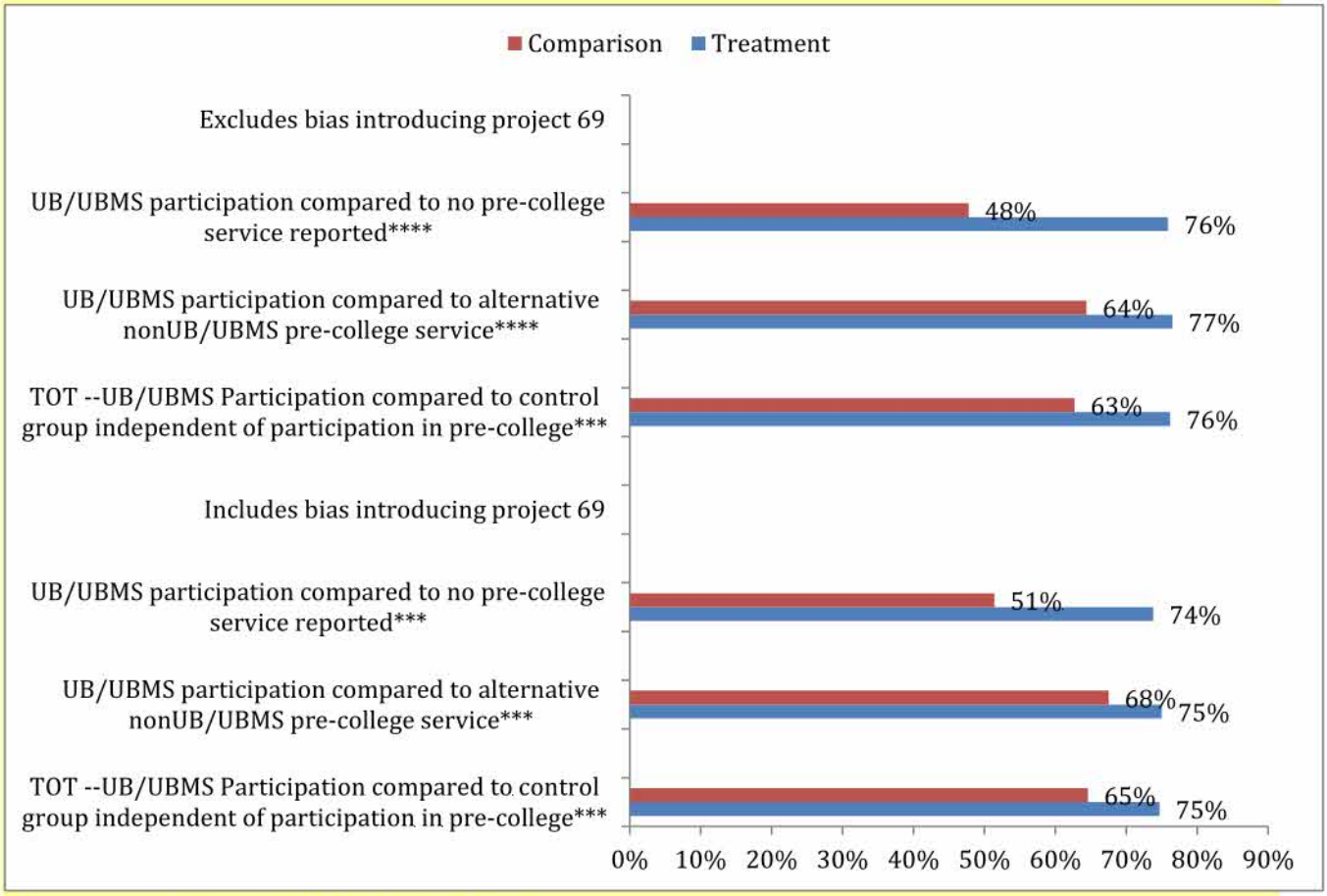
None of the Mathematica reports include comparisons of UB participants with those getting only other services or a serious consideration of the contamination issues related to the receipt of other similar but less intensive non-UB/non-UBMS services that were received by a majority of the control group. Cahalan reports that concern with the serious contamination issues, led one internal PPSS reviewer of the Mathematica Fifth Follow-up Report, Dr. Jay Noell, the PPSS UB Evaluation Technical Monitor/COR prior to Dr. Cahalan, to recommend that Mathematica acknowledge that the random assignment study had too high a level of contamination to be valid due to these similar but less intensive alternative services that had been provided often—precisely because the students did not get randomly chosen for Upward Bound. He recommended that the data be analyzed as a quasi-experimental design using instrumental variables regressions to model factors related to different levels of participation and to use these estimates in the second stage to control for the observed selection differences with regard to type of participation. Dr. Noell recommended that the ITT analyses be placed in an appendix, but not be used in assessing program effectiveness.

The *Cahalan Re-Analysis Report* presents the ITT and TOT analyses using models and methods similar to Mathematica except for standardization of outcomes and avoiding use of NSC for non-respondent enrollment and below BA degree estimation, but, following Dr. Noell's recommendation also includes some additional observational quasi-experimental design analysis using instrumental variables regression. These analyses compare outcomes for those who were in UB/UBMS with those who reported they participated in some other non-UB/UBMS pre-college supplemental service (see *Cahalan Re-analysis Report* chapter 4) and with those who reported not participating in any pre-college supplemental service program.

Exhibit B-14 gives results for postsecondary entrance in +4 years after expected high school graduation year and Exhibit B-15 gives results for award of BA degree in +6 years after expected high school graduation year. Both Exhibits show strong impacts for Upward Bound compared to both those who reported participation in some alternative less intensive pre-college service and compared to those who reported not participating in any pre-college services. For example, when **appropriate analyses controlling for selection bias are conducted, UB participants were 3.3 times more likely to obtain a BA in 6 years compared to those with no pre-college access services and 1.4 times as likely as those participating in only a less intensive service program such as Talent Search.** These strong positive results are not acknowledged in the reports. This failure to address this issue in the Mathematica reports has led to other researchers (see Haskins and Rouse 2013) to mistakenly

generalize from the erroneous Mathematica reports, that conclude that all federal college access programs have been shown not to be "effective". In fact 75 percent of the entire sample (treatment and control) had some form of supplemental pre-college services, and typically this was from another federally supported program.

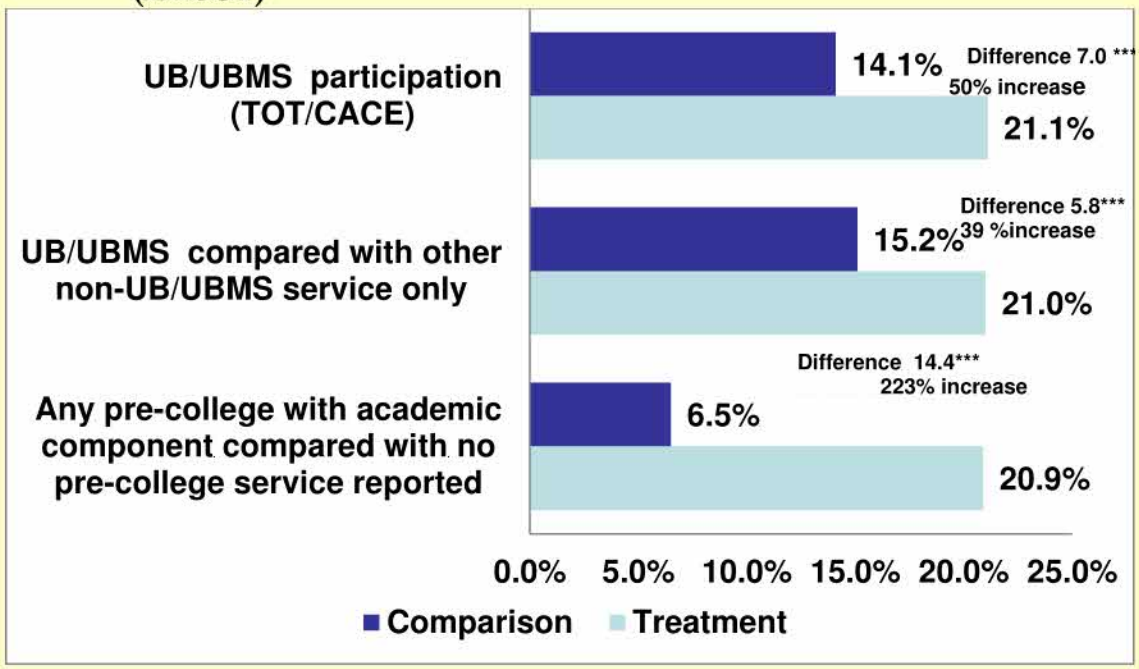
**Exhibit B-14: Estimates of relative impact of participation in various levels of pre-college access supplemental services on entry into postsecondary education within +4 years after expected high school graduation: National Evaluation of Upward Bound**




*/**/***/**** Significant at 0.10/0.05/.01/00 level.

NOTE: Based on 66 of 67 projects sampled. The estimates in the figures shown are based on longitudinal data over a 10- year period in an analysis using instrumental two-stage regressions that first model factors related to differences in participation in services and then use these factors in the second stage to control for participation selection bias factors.
SOURCE: Cahalan, Margaret: *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change?* The report can be accessed at the following site: http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml

**Exhibit B-15 . Instrumental variable Regression Results from the National Evaluation of Upward Bound for BA attainment in +6 years after Expected High School Graduation Year (EHSGY)**




*/**/***/**** Significant at 0.10/0.05/. 01/00 level.

NOTE: TOT = Treatment on the Treated (TOT); UB = Upward Bound; UBMS = Upward Bound Math Science. All estimates significant at the .01 level or higher. Estimates based on 66 of 67 projects in sample representing 74 percent of UB at the time of the study. One project removed due to introducing bias into estimates and representational issues. We use a 2-stage instrumental variables regression procedure to control for selection effects. SOURCE: Data tabulated January 2010 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education; study conducted 1992-9 to -2003-04.

**Violation 10**: **The Mathematica UB reports violate both the Joint Committee for Educational Evaluation Proprietary Standards that Stakeholders be fully informed concerning data that affects them and the AERA Standards related to the "sufficiency of the warrants and the transparency of the report"** *The Joint Committee Education Evaluation Standards state:*

> *P-6. The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results.*

This is among the most serious of the violations with the Mathematica reports. They are non-transparent in reporting positive impacts detected by the study and also in describing study issues—stating that the heavily weighted project 69 was typical of its stratum, and did not make a difference in conclusions. The reports also are not transparent in reporting the treatment-control group lack of balance on academic risk factors. The reports also state that the heavily weighted project 69 driving their no-impact

conclusions had "below average performance" when in fact the significant negative impacts found when this project is considered alone ---were due to the extreme uncontrolled differences between the treatment and control group in this project showing a failure of the random assignment procedures in this site. The reports are especially troubling because not only do they contain erroneous conclusions, but they also contain seemingly deliberate statements that mislead readers into thinking that the legitimate concerns, identified by PPSS internal and external reviewers, had been adequately addressed and did not make a difference in the study conclusions. **The QA re-analysis, however, conducted by Ed-PPSS technical monitoring staff found the assertions by Mathematica Policy Research of "no detectable impact" except for the award of certificates to be false.**

**Detailed Reports Documenting Standards Violations and Re-Analysis Results**

Detail on the major issues with the Mathematica reports and results of standards based re-analyses have been presented in several documents publically available at the addresses noted below.

- *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change? By Margaret Cahalan* a COE report published in 2009 and available at http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml.
- *The Council for Opportunity in Education (COE) Request for Correction submitted in 2012* is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf,
- *Expert Statement of Concern* with Regard to the Mathematica National Evaluation of Upward Bound signed by leading researchers can be found at http://www.coenet.us/files/ED-Statement_of_Concern_011712.pdf. The Statement accompanied the above referenced RCOE Request for Correction and was signed by researchers who had reviewed the request and found it cause of serious concern. The signers to the UB Evaluation Statement of Concern included the sitting presidents of the American Education Research Association (AERA) and the American Evaluation Association (AEA).
- Attachment D included with this package—Draft *Flawed Reports from the National Evaluation of Upward Bound Masked Significant and Substantial Positive Impacts: The Technical Monitors' Perspective* by Cahalan and Goodwin, forthcoming 2014

# April 7, 2014

**Margaret Cahalan**
Margaret.cahalan@pellinstitute.org
202-347-7430 ex 207
(b)(6) (c)

**David Goodwin**
(b)(6)

What Works Clearinghouse Quality Review Team
National Center for Education Evaluation and Regional Assistance
Institute of Education Sciences, U.S. Department of Education
555 New Jersey Avenue NW
Washington, DC 20208

Dear What Works Clearinghouse Quality Review Team,

We are writing to request that the WWC reconsider a rating of "meets evidence standards without qualification" that it has given to Mathematica's evaluation of the Upward Bound program. For reasons described below, we believe that there are major flaws in the evaluation design and analysis that the evaluator failed to acknowledge, resulting in incorrect impact estimates for the program's major outcomes. An independent analysis of the same data that attempts to address these flaws shows positive program impacts on college enrollment, application for financial aid and BA attainment. Since this evaluation has had a particularly large role in shaping policy debates and proposals, we believe that it is appropriate for WWC to reexamine its earlier rating in light of the evidence we present.

We also have concerns at the apparent conflict of interest given that, Mathematica was the contractor for the What Works Clearinghouse and that the ratings are included in the September 2009, *Practice Guide Helping Students Navigate the Path to College: What High Schools Can Do*, for which Mathematica staff are co-authors.

This request for rescinding the rating pertains to the following reports.

- Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular* ***Upward Bound****: Results from the third follow-up data collection*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbo ok
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do
- Seftor, N. S., Mamun, A., & Schirm, A. (2009). *The impacts of regular* ***Upward Bound*** *on Qpostsecondary outcomes 7–9 years after scheduled high school graduation*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbook
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

We are Dr. David Goodwin and Dr. Margaret Cahalan[1]. We served respectively as the first and last Contracting Officers Technical Representative (COTR) while serving within the Policy and Program Studies Services (PPSS) within the US Department of Education (ED). PPSS was the unit within ED responsible for technical oversight of the National Evaluation of Upward Bound.

During the final contract, after concerns about the study were raised, PPSS staff conducted a special Quality Assurance (QA) review and re-analysis of the data (PPSS QA review). This review found a number of unaddressed statistical and evaluation research accuracy and proprietary standards violations in the Mathematica Upward Bound reports that were serious enough to affect the validity and accuracy of the basic Mathematica conclusions from the study. The concerns raised by the ED PPSS monitoring staff concerning the Mathematica conclusions were made only after a long and careful review of the study methods and a complete set of data files. Files reviewed included the randomization file, the baseline survey and 5 follow-up surveys, 10 years of the Federal aid application and award files and 10 years of National Student

[1] Dr. Margaret Cahalan,is currently the Vice President for Research and Director of the Pell Institute for the Study of Opportunity in Higher Education of the Council for Opportunity in Education (COE). Dr. David Goodwin, who recently retired from the Gates Foundation is an Independent Consultant. He is the former PAS Division Director of the Policy and Program Studies Services (PPSS) of the US Department of Education (ED). Dr. Goodwin, served as the first Technical Monitor of the UB Evaluation, in the 1990s and at the time of the final contract served as Dr. Cahalan's supervisor. Dr. Cahalan, as a contractor served as the Project Director for several TRIO evaluations and performance reporting contracts, including the National Evaluation of Student Support Services and the Design Phase of the National Evaluation of Talent Search. After joining ED in late 2004, Dr. Cahalan supervised the COTR staff responsible for monitoring the final contract for the Mathematica UB evaluation and herself served as the Technical Monitor in the final months of the UB evaluation. In interest of full disclosure, Dr. Cahalan would like to note that she served as the Associate Director of the Survey and Information Services Division of the DC Mathematica Office from 1996 to 2002. During this time she supervised those persons in the Mathematica Survey Division responsible for the student follow-up surveys and transcript data collections and coding for the third and beginning of the fourth follow-up for the Mathematica Upward Bound evaluation. Her responsibilities did not include study design, analysis and reporting which were conducted under the Mathematica Research Division. During this time she was unaware of the major sampling and non-sampling error issues discussed in this document, however she did articulate concerns about issues of the 26 percent waiting list drop-outs, alternative service receipt by the control group, and survey non-response issues.

Clearinghouse (NSC) files. ED-PPSS also consulted with external experts who replicated the data re-analyses results. Although many of these statistical and evaluation research errors were identified prior to the final report's publication, departing political appointees decided to publish the report in early January of 2009.

In this letter we provide a summary of why we are registering this request for rescind of the rating of the Mathematica reports. This letter is accompanied by 4 attachments that provide key documentation of the material relevant to this request.

- **Attachment A**: Excerpt of UB Study Conclusions from the Executive Summary of the 2009 Mathematica Final Report (Seftor et.al 2009)
- **Attachment B:** Documentation of Key Standards Violations in the Mathematica Reports from the National Evaluation of Upward Bound (Prepared for this submission)
- **Attachment C:** Additional Documentation with Examples of Output from Logistic and Instrumental Variables Regression Models (taken from appendices B and D of Cahalan, 2009)
- **Attachment D:** *Flawed Contractor Reports from the National Evaluation of Upward Bound Masked Significant and Substantial Positive Impacts* (forthcoming, Cahalan and Goodwin, 2014)

As noted, Attachment A, to this letter is taken from the Executive Summary of the final (Sefter et.al. 2009) Fifth Follow-up report in which Mathematica lists major findings. Consistent with their earlier report published in 2004 (Myers et.al. 2004), Mathematica concluded and publically reported to Congress, OMB, and UB Stakeholders that Upward Bound did not have "detectable impacts" on the key legislative goals of the program. These impacts related to postsecondary entrance, application and award of financial aid, and attainment of postsecondary degrees or credentials. The only overall impact reported by Mathematica in their conclusions was a large impact on the award of *postsecondary certificates*.

When PPSS internal monitoring staff did a Quality Assurance (QA) review of the study, they found that the reports were based on a flawed sample design and flawed random assignment implementation. These issues were serious enough that all of the Mathematica impact estimates contained both representational error and a substantial systematic bias in favor of the control group. The analyses procedures followed by Mathematica also contributed to the erroneous conclusions because the contractor failed to standardize outcome measures for a sample that spanned 5 years of expected high school graduation. Mathematica also made improper use of the National Student Clearinghouse (NSC) data for survey non-respondents at a time when coverage was too low for enrollment and had not yet begun for degrees. The combined biases contained in the Mathematica impact estimates were serious enough to have resulted in a Type II statistical error of failure to detect positive impacts when they are present and the publication of erroneous evaluation conclusions concerning the Upward Bound program in both 2004 and 2009. The PPSS re-analysis work also found that when the identified error issues were addressed using standards-based statistical analyses that statistically significant and substantial positive impacts were observed for the Upward Bound program on key legislative goals of the program.

**Standards Used in ED-PPSS QA Review and in Mitigation Re-Analyses**

Listed below are the major education research standards and guidelines used and Exhibit 1 lists the specific standards and guidelines we believe are applicable to our concerns.

- *U.S. Department of Education Information Quality Guidelines (ED Guidelines)*
- *Joint Committee on Standards for Educational Evaluation (JCSEE).* http://www.jcsee.org/
- *National Center for Education Statistics (NCES) Statistical Standards*--- http://nces.ed.gov/statprog/
- *What Works Clearinghouse Standards (WWC)* --- http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1/
- *American Educational Research Association* (AERA) *Standards for Reporting on Empirical Social Science Research* in AERA Publications http://www.sagepub.com/upm-data/13127_Standards_from_AERA.pdf

Exhibit 2 identifies, and the pages to follow summarize, 10 interrelated violations of these standards in the Mathematica UB reports. These are the major reasons why, we, as persons who have carefully examined the data from this study, believe the above referenced reports should not be given the WWC rating of "meets evidence standards without reservations."

## Exhibit 1
## Key Information Quality Guidelines and Standards that are Applicable to the Concerns with Regard to the *Mathematica Upward Bound Reports*

**Department of Education Quality Information Guidelines**

Research and Evaluation information products should, at a minimum: …

- Pose the research or evaluation question in a balanced and unbiased manner;
- Provide an unbiased test of the question; …
- Present conclusions that are strongly supported by the data; ….
- Confirm and document the reliability of the data, and acknowledge any shortcomings or explicit errors in any data that is included;
- The source of data should be reliable. The sample should be drawn from a complete list of items to be tested or evaluated, and the appropriate respondents should be identified, *correctly sampled,* and queried
- Appropriate steps should be taken to *ensure that the respondents are a representative sample*;

**What Works Clearinghouse Handbook of Procedures and Standards**

A study may fail to meet WWC evidence standards if ..……

- It does not include a valid or reliable outcome measure, or does not provide adequate information to determine whether it uses an outcome that is valid or reliable. …..
- The intervention and comparison groups are not shown to be equivalent at baseline
- The overall attrition and or differential attrition rate exceeds WWC standards for an area.
- The measures of effect cannot be attributed solely to the intervention

……

**NCES Statistical Standards Concerning Non-Response and Coverage**

- **STANDARD 2-2-4:** A nonresponse bias analysis is *required* at any stage of a data collection with a unit response rate less than 85 percent. The extent of the analysis must reflect the magnitude of the nonresponse (see Standard 4-4).
- **STANDARD 3-1-2:** NCES data collections that are used as sampling frames for other NCES surveys must strive for coverage rates in excess of 95 percent overall and for each major stratum. **STANDARD 3-1-3:** If there is not evidence of a coverage rate of at least 85 percent of the target population, then frame enhancements such as frame supplementation or dual frame estimation must be incorporated into the survey study design.

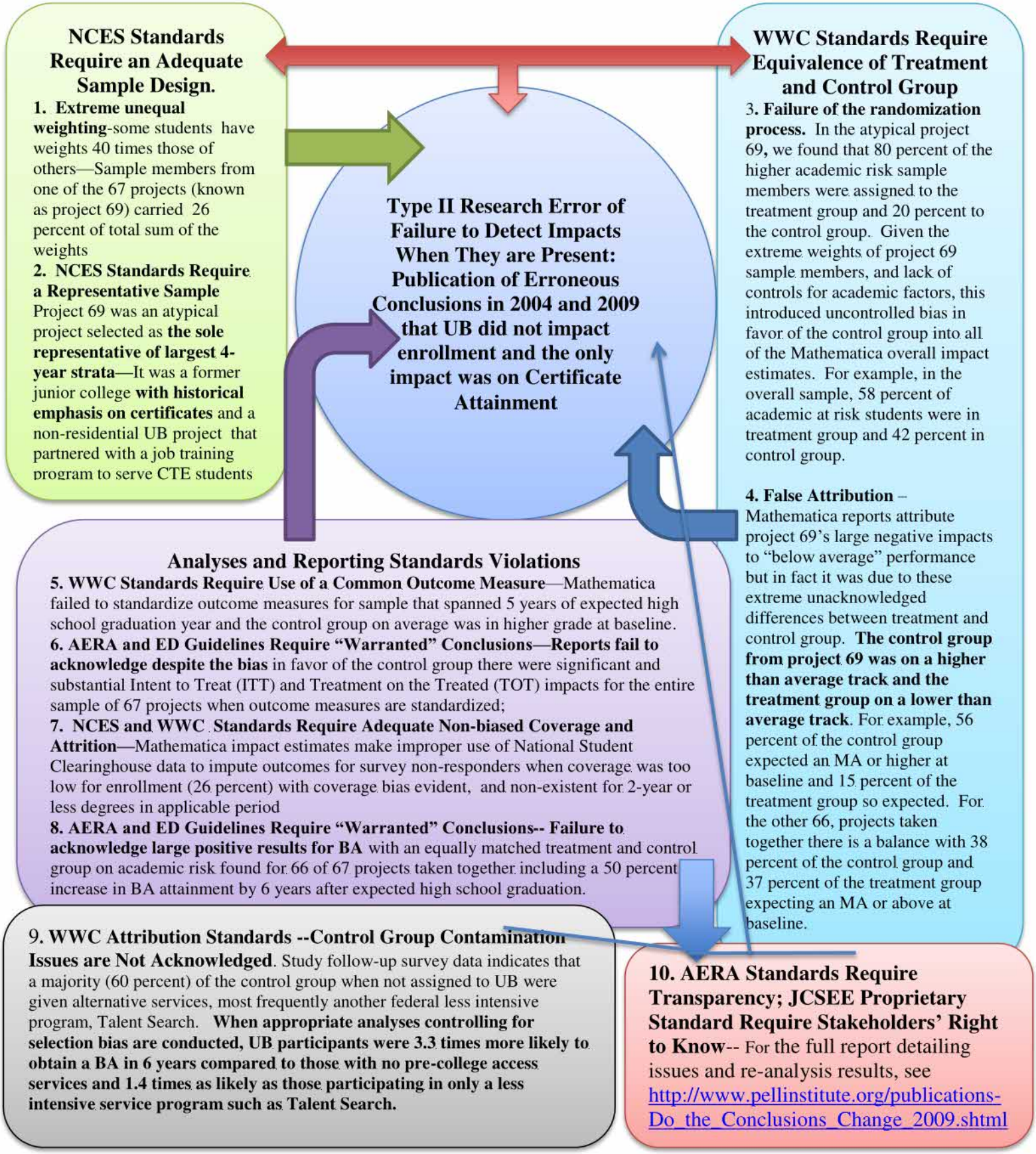
**Joint Committee on Standards for Education Evaluation Standards**: The Joint Committee Standards address ethics of research under the heading of Propriety. Standard P6 noted below discusses the full disclosure of findings

- **P6 Disclosure of Findings** The formal parties to an evaluation should ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results

**American Educational Research Association (AERA) Standards for Reporting on Empirical Social Science Research**

- Two overarching principles underlie the development of these reporting standards: the "sufficiency of the warrants" and the "transparency" of the report.

***Exhibit 2. NCES, WWC, JCSEE, AERA Standards and ED Guidelines, Violations in the Mathematica Upward Bound (UB) Evaluation Reports***



**Standards Violations in the Mathematica UB Reports**

The 10 violations identified in Exhibit 2 are discussed briefly in this letter and as noted documented in more detail in Attachments B to D included as separate attached files to this letter.

1. **ED Information Guidelines and NCES Standards Require an Adequate Sample Design —Serious Unequal Weighing.** Mathematica used a seriously flawed sample design to make inferences concerning the national average impact of Upward Bound with only one single project (known as project 69) selected to represent the largest study defined grantee 4-year and above public stratum. This design resulted in extreme unequal weighting in the final student level weighting stage. Some of the sample members from project 69 had weights that were 40 times those of the lowest weighted sample members and together project 69 sample members carried 26 percent of the sum of the weights (see Attachment B, Exhibit B-3).
2. **NCES Standards and ED Guidelines Require a Representative Sample for Estimation of Averages** In addition to the flawed design, the randomly selected project 69, was found to be a "bad draw" and "atypical" for its 4-year stratum. It did not possess the characteristics to be an accurate sole representative of the largest public 4-year BA and above granting stratum. Project 69 was a former junior college that historically awarded a large number of certificates. Its non-residential UB program was atypical for a 4-year UB grantee and partnered with a job training program. Adequate checks were not done to address the eligibility of project 69 to be the sole representative of the largest public BA and above set of UB grantees without introducing representational bias into the impact estimates. This fact combined with the extreme lack of balance in treatment and control group from project 69 (discussed below) led to a biased conclusion that the only positive impact of Upward Bound was on CTE certificates.
3. **WWC Standards Require A Balance Between the Treatment and Control Group At Baseline on Factors Likely to Impact Outcomes. (Non-Equivalence of Treatment and Control Group)** Due to a probable failure in the implementation of correct random assignment procedures in the project 69 site, there were also extreme differences between the heavily weighted treatment and control group in project 69 on academic factors, grade at entry into Upward Bound, and educational expectations. For example, in the highly weighted project 69, 80 percent of the students classified as higher academic risk were in the treatment group and 20 percent in the control group. Mathematica thus had a serious uncontrolled bias in favor of the control group on academic risk factors in all of their published overall impact estimates upon which they based their conclusions, For example, 58 percent of the academically at risk students were in the treatment group and 42 percent were in the control group. This lack of balance is not acknowledged in the Mathematica reports. (See Attachment B, Exhibits B-4 to B-7).

4. **WWC Standards Require that the Observed Impact be Attributable to the Intervention**. The Mathematica reports state that project 69 had "below average impacts" (negative impacts) and imply that this was because of "below average" project performance in this site. However, PPSS found in the QA review that in fact the so called "below average impacts" were due to the above noted severe lack of balance between the treatment and control group in project 69 and the failure of the random assignment implementation in this case to produce a balanced treatment and control group. As noted above, in the project 69 site there was observed an extreme lack of "balance at baseline" between the treatment and control group. The control group on average resembled applicants for an Upward Bound Math Science (UBMS) program being initiated at a nearby site, and not the typical applicants to the project 69 UB program with its CTE focus and partnership with a job training program. The control group was on average in a higher grade at baseline than the treatment group, and 56 percent reported expecting to obtain an MA or higher at baseline. Among the treatment group 15 percent expected an MA of higher at baseline and on average the treatment group resembled the less academically proficient students interested in CTE certificates and regularly served by the Project 69 UB grantee. Among the 66 other projects taken together 38 percent of the control group and 37 percent of the treatment group reported expecting an MA or higher at baseline. (See Attachment B, Exhibits B-4 to B-7)
5. **WWC Standards Require Use of a Common Outcome Measure for Impact Estimation. Lack of Precision and Standardized Common Outcome Measures.** In violation of the standard that common outcomes measures must be used, Mathematica used postsecondary outcome measures that were not standardized to expected high school graduation year. The sample spanned 5 years of expected high school graduation year cohorts. Hence the sample members had differences of up to 5 years in opportunity to enroll and complete postsecondary. This lack of precision also impacted the ability of the other variables used as controls in the regression models to function properly.(See Attachment B, Exhibits B-8 to B-10)
6. **AERA and ED Guidelines Require "Warranted" Conclusions.** The Mathematica reports fails to acknowledge and report statistically significant and substantial positive impacts estimates when standardization of outcome measures was implemented both with and without the bias introducing project 69 on postsecondary entrance, and application and award of financial aid These documented results were tabulated by PPSS technical monitors and conveyed to Mathematica in Spring of 2008, nine months before the Mathematica final report was published in early 2009. (See Attachment B, Exhibits B-8 to B-10)
7. **NCES Coverage Standards Require that the Data Sources Used Have Adequate and Non-Biased Coverage.** Mathematica made improper use of the National Student Clearinghouse data files for imputing enrollment and degree attainment for non-responders to the fifth follow up survey. In the most applicable period, NSC enrollment

coverage was estimated to be 26 percent and NSC had not yet even begun to collect degree or other credential information. Use of NSC data can only cautiously be used for BA receipt which would have occurred later after they had begun to collect degree information. There is also evidence of biased coverage in the NSC data due to the fact that Project 69 did not begin submitting enrollment data until after the most applicable period. **The Mathematica reports fail to acknowledge significant large positive impacts on award of any postsecondary credential based on survey data adjusted for non-response.** The improper NSC use led Mathematica to ignore their own positive impact estimates based on survey data adjusted for non-response that showed large significant impacts for award of any postsecondary credential by the end of the study period including a significant Intent To Treat (ITT) estimate of 13 percentage differences and a Treatment on the Treated (TOT) impact of 16 percentage difference (Seftor et.al. 2009, appendix tables C-7 and C14). Ignoring these positive and substantively meaningful impacts, in their highly publicized conclusions about Upward Bound, Mathematica reported finding no significant differences for award of any postsecondary degree—a key finding from the study. (See Attachment B, Exhibit B-11).

8. **AERA and ED Guidelines Require "Warranted" Conclusions-- Failure to acknowledge large positive results for BA for evenly matched treatment and control group.** Mathematica reported no impact on BA attainment, a major goal of the Upward Bound program. However, Mathematica failed to report the significant and substantial positive impacts (including a 50 percent TOT increase in BA attainment) that are observed for 66 of the 67 sampled projects. These 66 projects were found when taken together to have an equally balanced treatment and control group on academic risk factors and did not suffer from the serious representational issues of project 69 that were introduced into the overall sample given its extreme weights. (See Attachment B, Exhibit B-12)
9. **WWC Attribution Standards Specify that the Impact must be Attributable to the Intervention. --Control Group Contamination Issues are Not Acknowledged.** Mathematica fails to acknowledge the significance of the fact that a majority (60 percent) of the control group reported participation in an alternative supplemental pre-college service by the end of high school. Most frequently the alternative service into which those not selected for UB were placed was another less intensive pre-college federal TRIO program, Talent Search. Mathematica also fails to acknowledge that 26 percent of those randomly assigned to the treatment group were reported to have dropped off the so called "waiting list" due to student mobility by the time of random assignment to fill project openings. These cases were kept in the ITT treatment group in analyses although most did not enter UB nor have a realistic opportunity to do so.

To address these issues, in addition to replicating the Mathematica ITT and TOT impact analyses with outcome measures standardized to expected high school graduation year, the PPSS set of re-analyses also included using two stage instrumental variable regressions modeling selection effects to estimate the impact of participation in various

levels of services. These impact estimates indicated that participants in UB/UBMS[2] demonstrated significant and substantive positive impacts relative to those sample members participating only in an alternative typically less intensive program, such as Talent Search, and also relative to those not participating in any pre-college supplemental high school program. **For example, when appropriate analyses controlling for selection bias are conducted, UB participants were 3.3 times more likely to obtain a BA in 6 years compared to those with no pre-college access services and 1.4 times as likely as those participating in only a less intensive service program such as Talent Search.** None of these impacts are reported in the Mathematica reports. This lack of acknowledgement of control group contamination issues has led researchers such as Haskins and Rouse (2013), and also political budget offices, to mistakenly assume that the Mathematica (albeit erroneous) reports of "no impact" are indications that "all college access programs" are ineffective. In fact the UB evaluation was most definitely not a comparison of supplemental college access services vs. no services, as overall 76 percent of the total sample (treatment and control) reported participation in some form of supplemental pre-college services by the end of high school. (See Attachment B, Exhibits B-13 to B-15).

10. **The Mathematica UB reports lack transparency and violate the "stakeholders right to know" and the "sufficiency of the warrants" basic standards for evaluation research.** The Joint Committee for Standards for Educational Evaluation (JCSEE) Proprietary Standards specify that stakeholders must be fully informed concerning data that affects them and the AERA Standards stress the necessity of the "sufficiency of the warrants" and the "transparency of the report." The major study sampling and non-sampling error issues and positive impacts found by the PPSS monitoring staff in their QA re-analyses, are nowhere mentioned or acknowledged by the Mathematica reports. In fact the 2009 final report is written in such a manner as to give the impression that these issues are not of major concern.

**Detailed Reports Documenting Standards Violations and Re-Analysis Results**

Since the above referenced WWC Practice Guide was published in September of 2009, detail on the major issues with the Mathematica reports and results of standards based re-analyses have been presented in several documents publically available at the addresses noted below. In late 2009, Dr. Cahalan received permission from her supervisors to publish her results outside of the Department of Education and COE published her paper in October of 2009. In 2012, COE submitted a formal *Request for Correction* to the Department of Education, the contents of which are also available on the COE website.[3] The 2012 *Request for Correction* was accompanied by a *Statement of Concern* signed by leading researchers

[2] Upward Bound Math Science (UBMS) was a new initiative at the time of the start of the UB evaluation. Survey results indicated that about 12 to 14 percent of the control sample participated in Upward Bound Math Science a form of Upward Bound.

[3] The only response COE obtained from the 2012 Request for Correction was a letter from the OPEPD Assistant Secretary, indicating that ED would not re-consider the decision made by the departing Bush Administration in January 2009 to publish the report.

who had reviewed the correction request and found it cause for serious concern. The signers to the UB Evaluation *Statement of Concern* included the presidents at the time of the American Education Research Association (AERA), and the American Evaluation Association. We note that among the signers of the Statement of Concern is William Tierney, whom we have cc'd in this letter. Professor Tierney was the Chair Person of the panel for the Practice Guide in which the WWC UB ratings are presented.

- *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change? By Margaret Cahalan* a COE report published in 2009 and available at http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml.
- *The Council for Opportunity in Education (COE) Request for Correction submitted in 2012* is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf,
- *Expert Statement of Concern* with Regard to the Mathematica National Evaluation of Upward Bound can be found at http://www.coenet.us/files/ED-Statement_of_Concern_011712.pdf.
- Attachment D (included as an attachment to this letter) - *Flawed Reports from the National Evaluation of Upward Bound Masked Significant and Substantial Positive Impacts: The Technical Monitors' Perspective* by Cahalan and Goodwin, Forthcoming April, 2014

**Negative Consequences of Erroneous Mathematica Reports and WWC Ratings for Services for Low-Income and First Generation College Students**

These issues are not simply academic disputes with little consequences but are related to the basic judgments concerning the value of the work of the UB professional practitioners, and the grantee postsecondary institutions in seeking to foster the legislatively mandated goals of the program. As you may be aware, the results of this study have formed the basis for significant ED budget and other policy justifications for more than a decade. Based solely on the Mathematica UB study results, the Office of Management and Budget (OMB) rated the program as "ineffective." Then, justified by this rating and citing the 2004 Mathematica UB report findings, the Bush administration budget requests in FY2005 and FY2006 called for zero funding for all of the federal pre-college programs--Upward Bound, Upward Bound Math Science, Talent Search and GEAR UP. In November of 2011, the study report findings were reflected in the testimony to Congress of former Institute for Education Sciences (IES) Director, Grover T Whitehurst, asserting that federal programs such as Upward Bound had not been shown to be effective. More recently, in May of 2013, it has formed the justification for the mistaken assertion by a Brookings Policy Brief (Haskins and Rouse, 2013) that in general the federal college access programs "show no major effects on college enrollment or completion" and recommending that programs not able to demonstrate an effect should be defunded. These well-known authors state that their conclusions are based on the Mathematica Upward Bound study. They identify the Mathematica UB study as being the only evaluation of federal college access

programs to be given the highest study methods rating by the What Works Clearinghouse (WWC).[4]

We repeat our concern that is a very serious matter for the WWC to give a rating of "meets evidence standards without reservations" and to have this rating of the study be reported to Congress, the TRIO office, and academic and citizen stakeholders throughout the nation when there is clearly documented information that the Mathematica "no impact" conclusions concerning the Upward Bound program have been found to be erroneous. At this time we are respectfully writing to you to request your assistance in helping to address and publically correct this situation. This is not a new issue, but it is one that is long overdue for correction. Therefore we respectfully request that the What Works Clearinghouse (WWC) act in a timely manner to:

1. Publically rescind the WWC rating of the 2004 and 2009 UB reports as "Meets evidence standards without reservation" and correct the Practice Guide in which the UB reports were reviewed;
2. Consider the conflict of interest issues that the case of the Upward Bound evaluation raises concerning the behavior of the contractor for the study.

We call upon the persons responsible for this WWC rating to correct this endorsement of the 2004 and 2009 erroneous conclusions about UB program in a timely and public manner, before they can do more harm to the reputation of the WWC, the field of evaluation research, and most importantly to the availability of services for low-income and first generation students served by the TRIO and GEAR UP pre-college programs. We would be happy to meet with you at your earliest convenience should you wish to discuss any of the information concerning this matter.

Sincerely,

*Margaret Cahalan and David Goodwin*

Dr. Margaret Cahalan, Final ED-PPSS COTR for the Mathematica National Evaluation of Upward Bound, Retired SPCC Team Leader, PPSS, OPEPD, US Department of Education,
Current: Vice President for Research and Director Pell Institute for the Study of Opportunity in Higher Education, Council for Opportunity in Education (COE), Principal Investigator I-3 grant Using Data to Inform College Access Programming

Dr. David Goodwin, First ED-PPSS COTR for the Mathematica National Evaluation of Upward Bound; Retired Division Director, Policy Analysis Services (PAS), PPSS, OPEPD, U.S. Department of Education: Current: Independent Consultant, Gates Foundation.

---

[4] More recently, Mathematica President and CEO, Dr. Paul Decker, in his Nov 19, 2013 Presidential Address to the Association for Public Policy Analysis and Management (APPAM) presented Mathementic's erroneous impact estimate graphs as representing "the average impact of Upward Bound". These were based on data taken from the flawed 2009 Mathematica report (Sefter, et. al 2009) and were used to reaffirm publically that the UB evaluation study detected no impacts on major legislative goals. He characterized the response of what he called the "Youth Advocacy Community" to the Mathematica study as constituting "misdemeanors" and "felonies."

Margaret Cahalan

---

From:Margaret Cahalan
Sent:7 Apr 2014 17:44:13 +0000
To:Info@whatworks.ed.gov
Cc:(b)(6)
Subject:WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review
Attachments:SummaryletterUBWWCRatingConcernsCahalanGoodwinapril2014.pdf, Attachment A--Mathematic2009reportexcerptfromexecutivesummaryconclusions.pdf, AttachmentB_DocumentationofstandardsviolationsinUBreportsApril2014.pdf, AttachmentC_AdditonalUBConcernsDocumentationandSTATAoutput.pdf, AttachementD ForthcomingArticlecahalangoodwinapril2014.pdf
Dear What Works Clearinghouse,

We have attached 5 files providing our formal request for a Quality Review of the WWC ratings listed below.

- Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular **Upward Bound**: Results from the third follow-up data collection*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbo ok
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do
- Seftor, N. S., Mamun, A., & Schirm, A. (2009). *The impacts of regular **Upward Bound** on Qpostsecondary outcomes 7–9 years after scheduled high school graduation*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbook
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

The attached documents describe and document 10 specific concerns and violations of WWC Standards and of NCES and general statistical and evaluation research standards. We look forward to a timely and fair review of the concerns we have raised. If you have any questions regarding this submission, Dr. Goodwin and I would be happy to address these questions. Our contact information is below.

**Margaret Cahalan**
Margaret.cahalan@pellinstitute.org
202-347-7430 ex 207
(b)(6) (c)

**David Goodwin**

(b)(6)

Thank you for your consideration, Kind regards,
Maggie Cahalan

Margaret Cahalan, Ph.D, Vice President for Research, Director Pell Institute, email Margaret.cahalan@pellinstitute.org
Co-Principal Investigator i-3 grant Using Data to Inform College Access Programming
Council for Opportunity in Education (COE) www.coenet.us;
THE PELL INSTITUTE for the Study of Opportunity in Higher Education www.pellinstitute.org
1025 Vermont Ave., NW, Suite 1020, Washington, DC 20005, P: 202-347-7430 ex 207--C: (b)(6)

# *Attachment D*
# *Flawed Contractor Reports from the National Evaluation of Upward Bound Masked Significant and Substantial Positive Impacts (Forthcoming)*

**By Margaret Cahalan and David Goodwin**

## Executive Summary

In January 2009, in the last week of a departing Administration, the U.S. Department of Education (ED) published the final report in a long running National Evaluation of Upward Bound (UB) conducted by the contractor, Mathematica Policy Research. The report was published over objections from the Policy and Program Studies Services (PPSS) ED career technical staff who were assigned to monitor the final Mathematica contract. The report was also published after a "disapproval to publish" rating in the formal review process from the Office of Postsecondary Education (OPE), out of whose program allocation the evaluation was funded. The Mathematica reports from the UB study (Myers et. al. 2004; and Seftor et. al. 2009) have had a large impact on policy development for more than a decade, resulting in an OMB "ineffective rating" and zero budget funding justifications in President Bush's budget in FY2005 and FY 2006 for all of the federal pre-college programs, UB, Upward Bound Math Science (UBMS), Talent Search and GEAR UP.

**Reason for Speaking Out At this Time**. As the original (Dr. Goodwin) and final (Dr. Cahalan) Contracting Officers Technical Representatives (COTRs) for the study within the US Department of Education, our official job was to provide Technical Monitoring of the contracts. In the final of three sequential contracts, after concerns about the study were raised, we conducted a Quality Assurance Review (ED-PPSS QA review), and found that the impact estimations from the study being reported by the contractor were seriously flawed so much so that the basic conclusions Mathematica made concerning the efficacy of the Upward Bound program were impacted. While we have spoken out before on this topic, we are speaking out again in 2014, because of the on-going and recent citations of the erroneous findings from the study in Congressional testimony, policy briefs, and public speeches (Whitehurst, 2011, Haskins and Rouse 2013; Decker 2013). These erroneous findings continue to do serious reputational harm to the Upward Bound program.

**ED-PPSS QA Review**. The ED-PPSS QA review involved an internal review and analysis of all data files from the study, as well as consultation and replication of results by external statistical experts. The data files reviewed included: the initial sampling frame, the baseline survey, five follow-up surveys, student transcripts, 10 years of federal aid files and 10 years of National Student Clearinghouse (NSC) data. The ED-PPSS QA found that the Mathematica reports were seriously flawed, made unwarranted conclusions about the Upward Bound program and were not transparent in reporting. Moreover statistically significant and educationally meaningful positive impacts on the key legislative goals of the Upward Bound program were clearly found when the study errors were addressed using standards based statistical methods. These positive impacts are unacknowledged in the Mathematica reports. Below are highlights from the PPSS QA review and re-analysis.

**Major Flaws Identified in the Reports**. Major statistical and evaluation research standards violations were found including: 1) A flawed sample design with severe unequal weighting in which the highest weighted students had weights 40 times those of the lowest weighted students and one single project of 67 carried fully 26 percent of the weight; 2) Serious representational errors with one single atypical former 2-year college with an historical focus on certificates selected to represent the largest 4-year and above degree granting stratum; 3) Severe non-equivalency of the treatment and control group on academic risk,

grade at entrance, and educational expectations leading to uncontrolled bias in favor of the control group in all of the impact estimates upon which conclusions were made; 4) Failure to use a common standardized outcome measures for a sample that spanned 5 years of expected high school graduation year; 5) Improper use of National Student Clearinghouse (NSC) data to impute survey non-responders' enrollment and degree attainment status when coverage was far too low and non-existent for 2-year and below degrees, with bias clearly evident; 6) False attribution of large negative impacts in the project with extreme weights to "poor performance" ignoring the extreme bias in favor of the control-group in this project's sample ; 7) Lack of addressing issues of control group receipt of alternative but less intensive federal pre-college services received by the majority (60 percent) of the control group members; and 8) Lack of reporting transparency and failure to acknowledge strong positive impacts of UB on key program goals that are found when these errors are addressed using standards based statistical and evaluation research methods.

**PPSS Re-Analysis Found Strong Positive Impacts**. Contrary to the Mathematica conclusions that the only overall impact was on certificate attainment, the ED-PPSS QA re-analysis conducted by ED internal monitoring staff found that when NCES and What Works Clearinghouse (WWC) standards were followed to mitigate or correct the errors noted above, there were statistically significant and substantively meaningful positive results for the Upward Bound program. These impacts were on the major legislatively-mandated goals of the program----postsecondary entrance, application for and award of financial aid, and degree attainment (see Figures 5 to 8). The impacts included a 50 percent Treatment on the Treated (TOT) increase in BA degree attainment within six years of expected high school graduation using the balanced treatment and control group (Figure 6). Instrumental variables regression controlling for selection factors revealed that 75 percent of UB/UBMS participants entered postsecondary within one year of high school graduation compared to 62 percent of those who received only a less intensive service such as Talent Search, and 45 percent of those who reported no pre-college service receipt (figure 7). PPSS also found that UB/UBMS participants were 3.3 times more likely to obtain a BA in six years when compared to those reporting no participation in college access supplemental services and 1.4 times as likely when compared to those who reported participating in less intensive supplemental services (Figure 8). For the full re-analysis report detailing issues and full documentation of the re-analysis results, see http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml

**Support for "*COE 2012 Request for Correction*" Submitted to ED in 2012 and for the "*2014 Request to Rescind*" the WWC UB Study Rating** The article concludes that the non-transparent published reports from the National Evaluation of Upward Bound suffer from what is known as a Type II study error, or a failure to detect positive impacts when they are present. Thus the Mathematica conclusions that UB had no impact on postsecondary entrance, financial aid or degree attainment outcomes except for a positive impact on the award of certificates are incorrect. The article expresses support for the Council for Opportunity in Education (COE) formal *Request for Correction* submitted to the Department of Education in 2012 calling for the Mathematica reports to be corrected or withdrawn. The article also supports the 2014 request that the What Works Clearinghouse (WWC) "rescind" the 2009 rating given to the UB study reports of "meets evidence standards without reservations." The 2012 request was accompanied by a Statement of Concern signed by leading researchers in the field, including the sitting presidents of the American Education Research Association (AERA) and the American Evaluation Association (AEA). The complete text of the *Request for Correction* is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf, and the *Statement of Concern* signed by leading researchers can be found at http://www.coenet.us/files/ED-Statement_of_Concern_011712.pdf.

## Introduction

In January 2009, in the last week of a departing Administration, the U.S. Department of Education (ED) published the fourth and final report in a long running National Evaluation of Upward Bound (UB) (Myers and Schirm 1996; 1999; Myers et. al. 2004; and Seftor et. al. 2009). The 2009 report was published over the objections of the ED career technical staff assigned to monitor the final contract, and after a "disapproval to publish" rating in the formal review process from the Office of Postsecondary Education (OPE), out of whose program allocation the evaluation was funded.

*Dr. Cahalan is Vice President for Research and Director of the Pell Institute for the Study of Opportunity in Higher Education of the Council on Opportunity in Education (COE). Dr. Cahalan supervised the staff serving as the UB evaluation's technical monitors and served in this capacity herself in the final few months of the UB evaluation. She is currently the Co-PI of the COE i-3 project "Using Data to Inform College Access Programming."*

*Dr. Goodwin is currently an independent consultant for the Gates Foundation. He is the former Director of the unit within the U.S. Department of Education responsible for the UB Evaluation. Dr. Goodwin was Dr. Cahalan's supervisor at the time of the final Mathematica UB Contract. He was the UB study monitor when the study was first begun in 1992.*

Upward Bound (UB) is a Federal program, begun in 1965, designed to provide college readiness through supplemental academic services, as well as college awareness, leadership, and counseling services. Congressionally-mandated eligibility requirements specify that two-thirds of the high school participants must be low-income (defined as 150 percent of the poverty level) and students who would potentially be the first person in their family to obtain a bachelor's (BA) degree (known as "first-generation college" students). The other one-third must be either low-income or first-generation**.** Upward Bound is one of the first and considered a model flagship Federal program. It is also one of the more intensive low-income and first-generation college access programs with an average cost per student of about $4,300. There are about 900 Upward Bound (UB) and Upward Bound Math Science (UBMS) programs across the country. Project grantees responsible for implementing UB are 4-year and 2-year postsecondary institution and community organization grantees who together serve about 65,000 high school students yearly. The program has a strong academic focus with an intensive six-week summer traditionally residential program that is held on a college campus followed by weekly academic year sessions throughout high school. As specified in the authorizing legislation, all Upward Bound projects must provide instruction in mathematics through pre-calculus, laboratory science, foreign language, composition and literature through summer programs on a college campus and academic year supplemental services. The goal of Upward Bound is to increase the rate at which low-income and potentially first-generation college participants complete secondary education and enroll in and graduate from institutions of postsecondary education. UB and UBMS grantees hold competitive five-year grants to administer UB services to low-income and first-generation students in high-needs target high schools in their local communities.

The random assignment longitudinal study followed low-income and "potentially first-generation-college" students from middle school or early high school through six to 10 years after their expected high school graduation year (EHSGY). The study was conducted under a

series of three contracts with a baseline and five follow-up student surveys by Mathematica Policy Research (Mathematica).

The results of this seemingly high-quality random assignment study have formed the basis for significant policy justifications—most notably a Bush administration budget request to eliminate funding for Upward Bound and other federal pre-college access programs—Talent Search and GEAR UP, and a decision by the Office of Management and Budget (OMB) to rate the program as "ineffective." In November 2011, the study report findings were reflected in the testimony to Congress of former Institute for Education Sciences (IES) Director Grover T Whitehurst, asserting that federal programs such as Upward Bound and Head Start had not been shown to be effective. More recently, in May 2013, it has formed the justification for the assertion by a Brookings Policy Brief (Haskins and Rouse, 2013) that in general federal college access programs "show no major effects on college enrollment or completion." These well-known authors state that their conclusions are based primarily on the Mathematica Upward Bound study. They identify the Mathematica UB study as being the only evaluation of federal college access programs to be given the highest study methods rating by the What Works Clearinghouse (WWC), a clearinghouse, co-incidentally also run under an ED contract to Mathematica.

Ironically, as Technical Monitors for the evaluation while working at ED-PPSS, we found in a Quality Assurance (QA) review of study design and data files that the widely-cited reports from this evaluation were not transparent and made unwarranted conclusions concerning the Upward Bound program. We concluded that the Mathematica reports were seriously flawed in terms of statistical sampling standards violations and importantly had a serious uncontrolled statistical bias in favor of the control group on academic risk factors. These identified biases violate a basic National Center for Education Statistics (NCES) and WWC Standards that the sample be representative of the population of interest and that the treatment and control group be balanced and equivalent on baseline factors related to outcomes. Importantly, we also found, when we conducted a re-analysis based on NCES and WWC standards and the recommendations of independent external statistical reviewers, that there were statistically significant and substantively strong positive results for the Upward Bound program. These impacts were on the major legislatively-mandated goals of the program----postsecondary entrance, application for and award of financial aid, and attainment of bachelors' (BA) degrees and other postsecondary degrees or credentials. We concluded that the non-transparent published reports from the National Evaluation of Upward Bound suffer from what is known as a Type II study error, or a failure to detect positive impacts when they are present.

We made our concerns and the QA re-analysis positive results well known to Mathematica and the Department of Education at the time (Cahalan 2009). As the ED Technical Monitors for the study, we reiterate our serious concerns publicly now in the light of repeated use of the flawed Mathematica results in Congressional testimony, policy briefs, and public speeches (Whitehurst, 2011, Haskins and Rouse 2013; Decker 2013). We also do so in order to support the formal *COE 2012 Request for Correction* of the Mathematica final report, submitted to ED almost two years ago, by COE and their affiliated regional Educational Opportunity Organizations. These organizations represent TRIO program stakeholders in the evaluation. The COE request for correction was accompanied by a *Statement of Concern* signed by, among others, the Presidents of the American Evaluation Association (AEA) and the American Education Research

Association (AERA). Each of the signers of the Statement of Concern had reviewed the *COE Request for Correction* prior to signing the *Statement of Concern*. We are also writing this report in order to support a formal *Request to Rescind* the rating given by the What Works Clearinghouse (WWC) of "Meets evidence standards without reservations" given to Mathematica Upward Bound reports in the 2009, WWC Practice Guide entitled: Helping Students Navigate the Path to College: What High Schools Can Do.

Before discussing our QA findings in more detail, we wish to make clear that this article is not intended to be a general critique of the random assignment method nor a post-hoc effort to "fish" for positive study findings. Nor is the article intended to discredit the study as a whole. While we object strongly to the failure of Mathematica to address the flaws in their impact estimates or to acknowledge the positive results obtained when these issues are addressed using standards based methods, we also believe that the *National Evaluation of Upward Bound*, when corrected for sampling and non-sampling error, can be a very useful and informative study in the area of pre-college research. The essence of our findings is detailed below.

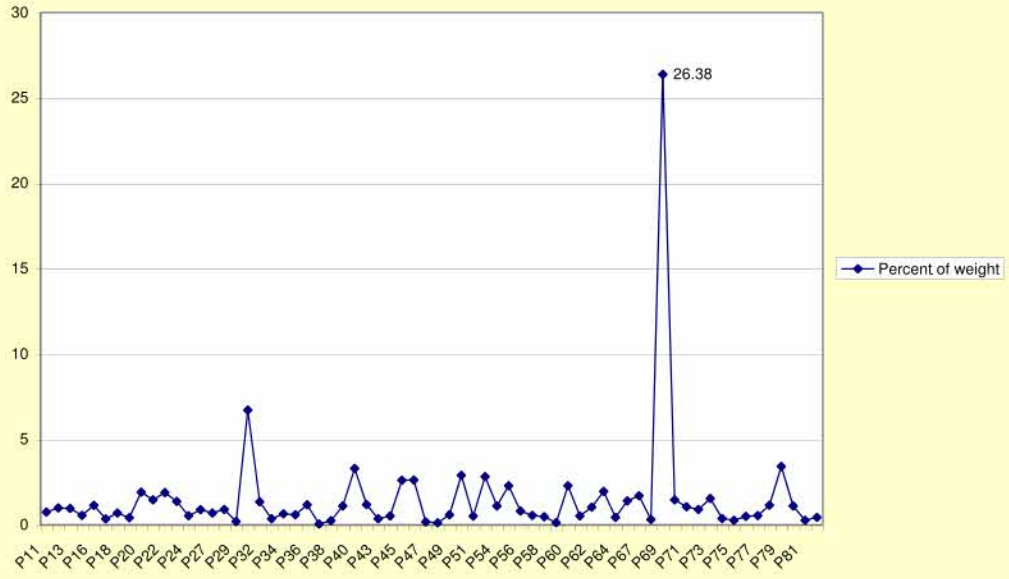
**Major Findings from the Technical Monitors Review of Study Issues**

**Seriously Flawed Sample Design and Severe Unequal Weighting.** The design for this study was unusual and overly ambitious and unfortunately resulted in a multi-stage sample with one project carrying 26.4 percent of the final student weights. In what reviewers have called a "seriously flawed sample design" that does not meet NCES standards, only one project in the sample (called project 69) was selected to represent the largest study defined 4-year and above public grantee stratum. Furthermore, because of an unusually large number of "baseline" surveys from interested students submitted by project 69, in the final stage of weighting, project 69 carried fully 26 percent of the weights. Figure 1 shows just how extreme the unequal weighting was from project 69. The method of counting baseline surveys submitted by the sampled projects as "applicants" and constituting a so called "waiting list" and then weighting to the number of baseline surveys (considered applicants) within project defined sub-strata further confounded the already-flawed first stage sample design. In addition, projects used different recruitment methods to obtain the "waiting list" based on returned baseline surveys and were allowed to create project specific sub-strata from which students were randomly selected at differential rates. Subsequently there were large differences among the sampled projects in the ratio of baseline surveys submitted to the number of project openings over the period. The weights were the inverse of the probability of selection at each of the stages (project and student applicant level). Because project 69 was supposedly representing a very large number of both projects and applicants, this flawed design meant that the outcomes of some students from the project 69 "waiting list" carried a weights that were 40 times those of the lowest weighted students (for example, some project 69 sample members had weights of 158 while the lowest

***In what reviewers have called a "seriously flawed sample design" that does not meet NCES standards, only one project in the sample (called project 69) was selected to represent the largest study defined 4-year public stratum and carried fully 26.4 percent of the weight.***

weighted sample member among all the projects carried a weight of 4). Mathematica reports that were published over almost a 10 year period did not reveal these serious sample design issues.

**Figure 1. Percentage distribution of sum of the weights of the 67 projects making up the study sample: National Evaluation of Upward Bound, study conducted 1992-93 to2003-04**




**NOTE**: Of the 67 projects making up the UB sample just over half (54 percent) have less than 1 percent of the weights each and one project (69) accounts for 26.4 percent of the weights.

**SOURCE:** Data tabulated December 2007 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Planning Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education,: study conducted 1992-93 to 2003-04.

**Atypical Project Selected As Sole Representative of Largest Stratum.** Unfortunately, project 69, whose students carried 26 percent of the weight, was also found to be atypical. Randomly chosen as the sole representative of the largest study defined 4-year and above grantee stratum, the project 69 grantee institution had historically been a junior college, offering associate and certificate programs taken over to serve as a branch of a nearby 4-year city-wide college system. Project 69's UB program was non-residential and partnered with a job training program serving Career and Technical Education (CTE) target minority high schools. It thus had a higher-than-average, especially for a 4-year grantee, percentage of its UB participants who were interested in seeking less than 2-year vocational certificates.

***The ED staff QA review found that project 69 was "atypical" of the 4-year stratum for which it was the sole representative.***

The study reports do not reveal project 69's representational issues, and indeed Mathematica's final report specifically asserts that project 69 is an adequate sole representative of the types of projects likely to be present within this, the largest 4-year and above study stratum (Sheftor, et. al. 2009). The stratum project 69 was supposedly representing and that justified its

26 percent weight was a large combined stratum of average sized projects housed at 4-year colleges and universities. It included the major flagship research universities as well as small 4-year liberal arts colleges that had UB grants at the time. Neither of these types of 4-year and above grantees could be adequately represented by project 69.

**Serious Lack of Balance between the Treatment and Control Group.** A basic standard of the What Works Clearinghouse and random assignment studies generally is that in order to make valid impact estimates, the treatment and control group must be equivalent at baseline on factors related to outcomes. Although the random assignment method is intended to ensure that treatment and control groups are equivalent (and did so quite well for the combined UB sample without project 69), in project 69, the QA review found major differences between the treatment and control groups on factors related to outcomes. The imbalance in project 69 was so large that some external reviewers reported they suspected a failure to implement the random assignment correctly in this project. For example as shown in Figure 2 below, 80 percent of the academically at-risk students from the project 69 sample were in the treatment group (randomly assigned to Upward Bound in middle or early high school), while 20 percent of the academically at-risk students were in the control group (not randomly assigned to UB in middle or early high school).

**Figure 2 Project 69 has severe imbalance in favor of control group: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-2004**

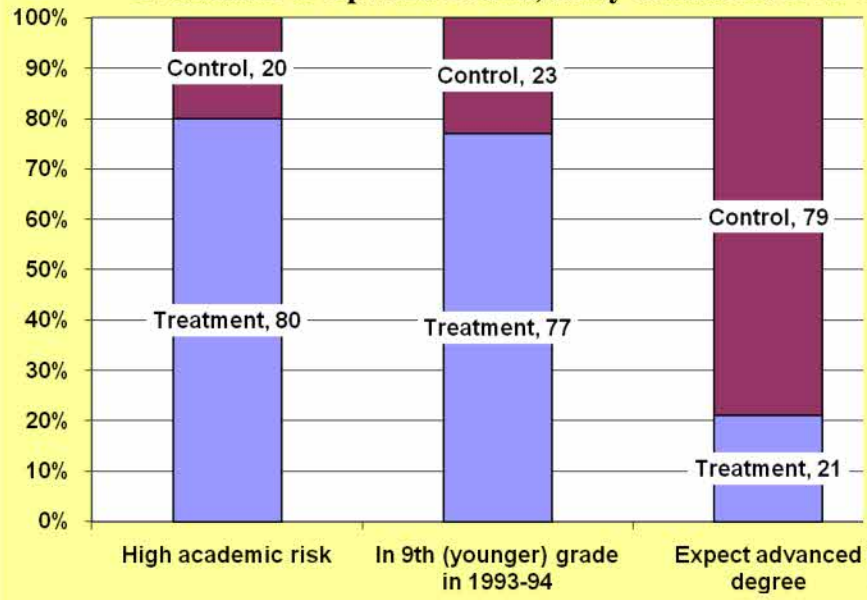



Figure 2 reads, 80 percent of the high academic risk students were in the treatment group and 20 percent in the control group; 79 percent of those expecting to obtain an advanced degree MA or higher were in the control group and 21 percent in the treatment group. This indicates a severe lack of balance between the treatment and control group

For project 69, the treatment sample on average resembled the vocational programming emphasis of the project, with a larger than average for a 4-year grantee of participants interested in certificate programs; while the control group on average resembled the typical Upward Bound Math Science (UBMS) applicant with a larger percentage on average interested in obtaining advanced degrees (56 percent). After the identity of project 69 became known to ED at the end of the final contract, in researching the project 69 issue, we found that there was a neighboring newly formed UBMS project operating in the region. As seen in Figure 2, the control group members were in a higher grade, were more academically proficient, and had considerably higher educational expectations at baseline. This suggests that the unusually large number of baseline surveys (n=85) collected by project 69 relative to their actual openings may have been because they included those students who were actually applying for the neighboring UBMS program from a high school science and technology magnet program also located at one of the project 69 target schools along with the Vocational Career and Technical Education program. As Technical Monitors, we discovered these issues only gradually when we did direct QA analysis of the data files to discover why project 69's Upward Bound program had demonstrated such seemingly negative impacts on postsecondary outcomes relative to its control group.

***The UB study analyses violate the basic What Works Clearinghouse standard that the treatment and control group be equivalent on baseline factors related to outcomes.***

Unfortunately, the severe non-equivalency in project 69 combined with the extremely large weights for the students from this project resulted in an imbalance in the overall sample and an uncontrolled bias in favor of the control group in all of the Mathematica impact estimates (Mathematica had no controls for academic risk factors in their analysis). For example, in the overall sample with project 69 included, 58 percent of the academically at-risk students were in the treatment group and 42 percent in the control group (Figure 3). In contrast, when we did balance checks on the combined sample without project 69, we observed a good balance between the treatment and control group on these same factors, with for example, 51 percent of the academically at-risk students in the treatment group and 49 percent in the control group (Figure 4).

**Figure 3. Imbalance in Overall Upward Bound Sample with Project 69 included: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-2004**

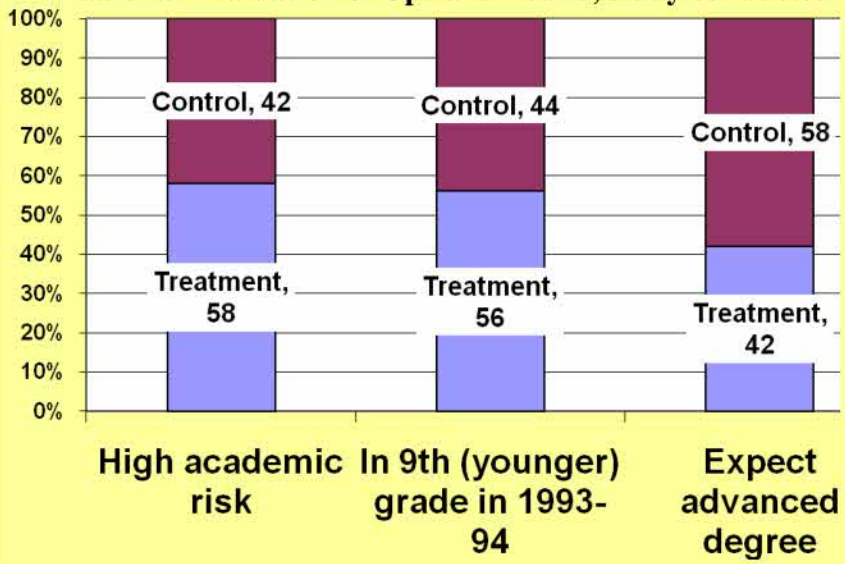



Figure 3 reads, for example: In the overall sample, among the high academic risk students,58 percent were in the treatment group and 42 percent in the control group

**Figure 4. More Balanced Treatment and Control Group for 66 other projects taken together: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-2004**

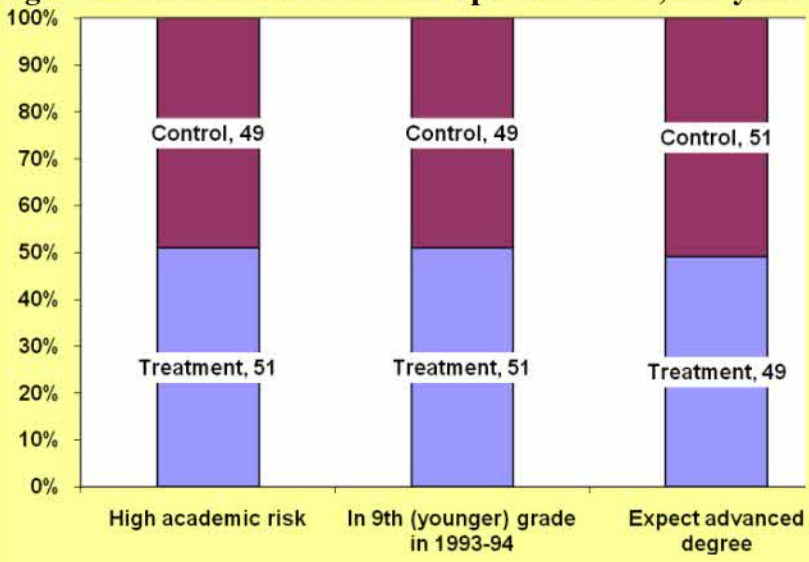



Figure 4 shows the balance between the treatment and control group on key factors when project 69 is excluded

.

**Lack of Standardization of Outcome Measures to Expected High School Graduation for a Sample that Spanned Five years of Expected High School Graduation Year**. The issues noted above were aggravated by the fact that Mathematica, in violation of the NCES and What Works Clearinghouse standards, did not standardize the outcome measures for a sample that spanned five years of expected high school graduation years. Mathematica argued that randomization made this unnecessary. However, balance checks done by ED monitoring staff found that on average, the control group was in a higher grade in a fixed academic year than the treatment group (see Figure 3). In addition, to the obvious issues related to differences in levels of potential opportunity to enter postsecondary and complete degrees over five years of expected high school graduation years, this lack of standardization also confounded the ability of the other variables in the regression models to function in a meaningful way to control for baseline differences.

***The Mathematica reports, use unstandardized outcome measures for a sample that spanned 5 years of expected high school graduation dates violating NCES and What Works Clearinghouse standards requiring use of common standardized outcome measures.***

**Improper Use of National Student Clearinghouse (NSC) Data.** In violation of NCES standards, the final report of the Mathematica study also makes improper use of NSC data for imputation outcome measures for survey non-responders in a very early period when enrollment coverage was too low and when degree coverage for 2-year and less than 2-year degrees was nonexistent. This improper use of NSC introduced bias into the conclusions Mathematica reported for the study. For example, Mathematica ignored their own significant and substantial positive impact results based on fifth follow-up survey data adjusted for non-response for the award of "any postsecondary degree or credential" (Seftor et.al. 2009, see appendix C), and falsely reported that the study only detected postsecondary credential impacts only for award of certificates.

Mathematica's own estimate of attainment of "any postsecondary degree or credential" based on responders to the fifth-follow-up survey adjusted for non-response shows a positive substantial and significant Intent To Treat (ITT) impact of UB on award of "Any postsecondary degree or credential" of 13 percentage points (55 percent for UB and 42 percent for the control group) and a Treatment On the Treated (TOT) estimate of a 16 percentage point difference—(Seftor et. al. 2009 Appendix tables C-7 and C14). Against ED Technical Monitors' recommendation and that of the IES external reviewers to be conservative in use of NSC, Mathematica chose to present in the text tables in the body of the report and base their conclusions only those estimates that used NSC data for non-responders to the fifth follow-up—coding the 25 percent of the sample who were survey non-responders and who were not found in NSC as "not having any degree or certificate." The significant and large positive results noted above, tabulated by Mathematica itself, are included in Mathematica's appendix tables C-7 and C-14 but not mentioned in the text body presentation of study conclusions.

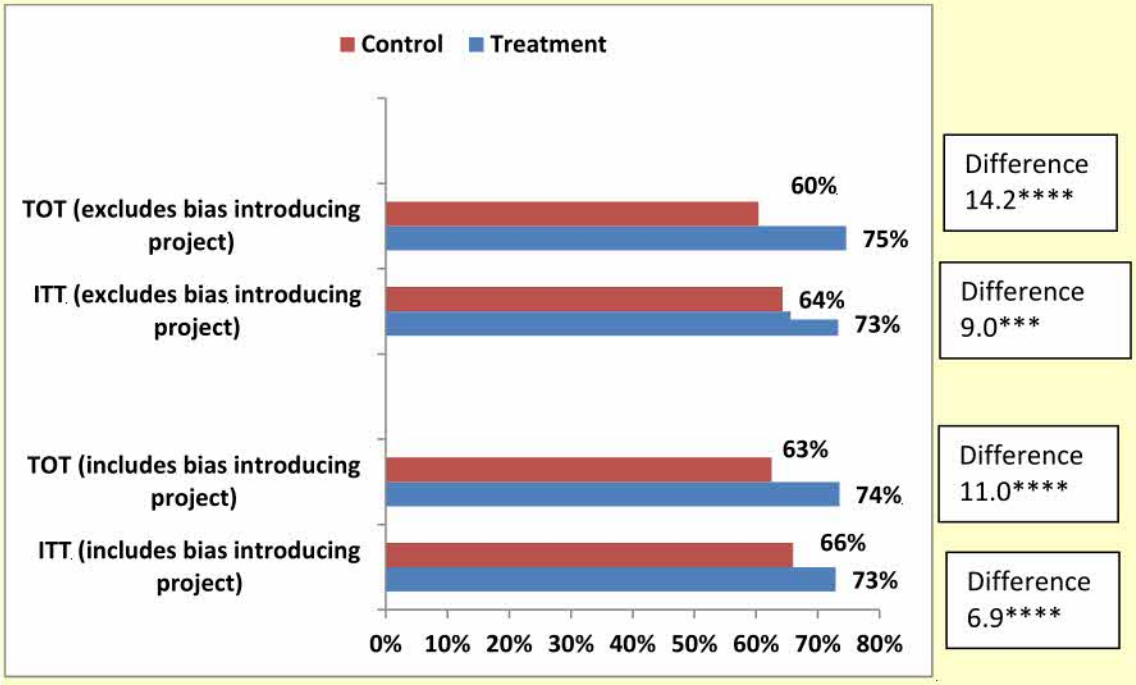
**Alternative Re-Analysis Major Findings**

As the issues within the Mathematica UB reports became known to ED staff, we began to consult outside experts and to use NCES and WWC Standards as guides to mitigate the issues. We prepared impact estimates that we considered more robust containing less statistical bias. In conducting the re-analysis, we standardized outcome measures to expected high school graduation year. To maximize response, the re-analyses also included information from each of the three applicable follow up surveys (third through fifth), and used 10 years of federal aid and award files to supplement the survey data. However, following NCES standards, we avoided use of the NSC for enrollment and degrees less than the BA due to lack of coverage in this early period in the NSC history. Following expert advice, we prepared and reported all impact estimates with and without project 69 and included impact estimates for the sample, weighted and unweighted. For the full re-analysis report detailing issues and full documentation of the re-analysis results see http://www.coenet.us/files/files- do_the_Conclusions_Change_2009.pdf ).

***The ED re-analysis standardized outcome measures and found positive outcomes with and without project 69 on enrollment and award of financial aid***

**Positive Impacts on Postsecondary Entrance and Financial Aid With and Without Project 69.** The QA re-analysis of the data standardizing outcome measures to expected high school graduation year (EHSGY) found there were substantial and statistically significant positive impacts on postsecondary entrance, application and award of financial aid, and completion of any postsecondary degree or credential with and without project 69. Figure 5 gives an example of these findings for postsecondary entrance after 1 year. Similar impacts were seen for enrollment four years after expected high school graduation year.

**Figure 5. Treated on the Treated (TOT) and Intent to Treat (ITT) estimates of impact of Upward Bound (UB) on postsecondary entrance within +1 year (18 months) of expected high school graduation year (EHSGY) 1992-93 to 2003-04**




*/**/***/**** Significant at 0.10/0.05/. 01/00 level.

**NOTE**. Model based estimates based on STATA logistic and instrumental variables regression and also taking into account the complex sample design. Based on responses to three follow-up surveys and federal student aid files.

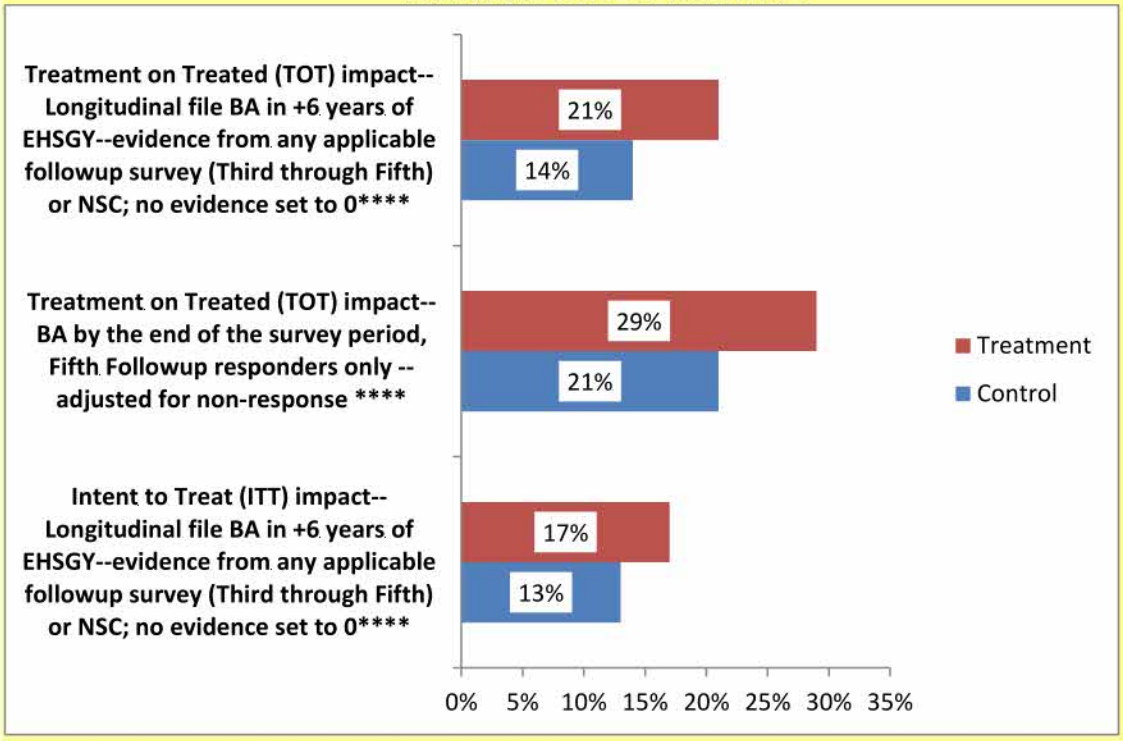
**SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, and federal Student Financial Aid (SFA) files 1994-95 to 2003-04. (Excerpted from the *Cahalan Re-Analysis Report*, Figure IV)

**BA Attainment Impact Analysis.** As noted the representational issues combined with the treatment control group non-equivalency in the heavily weighted project 69 introduced a serious uncontrolled bias into the Mathematica impact estimates. This was especially apparent for BA receipt and could not be addressed adequately by simply standardizing outcomes to expected high school graduation. As noted on average the control group from project 69 resembled Upward Bound Math Science program applicants, being in 10th grade at application, having advanced degree expectations and being more academically proficient. In contrast the treatment group from project 69 on average was comprised of students interested in obtaining certificates,

***Among the most impressive of the re-analysis findings was that when the treatment and control group are equivalent, there was a 50 percent increase in BA attainment by 6 years after expected high school graduation date for those students randomly assigned to UB and who participated in the program***

more academically at-risk and having lower expectations. In fact, the project 69 treatment group was contributing fully one-third of the weights for those designated as academically at-risk in the overall sample. The PPSS external advisor, Dr. Chromy, recommended basing the BA analysis on the 66 projects that together have a balanced treatment and control group and acknowledging that the study cannot adequately represent the large 4-year and above grantee stratum for which project 69 is the sole representative. The QA re-analysis found that when there is an equivalent baseline treatment and control group as is present when 66 of the 67 projects are taken together, there are also strong positive impacts on BA attainment. As seen in Figure 6, the Treatment on the Treated (TOT) impact analyses revealed that those sampled students randomly assigned to UB and/or who participated in the program had about a 50 percent increase in likelihood of obtaining a BA in six years compared with those not randomly assigned and who did not participate in the program. The Intent to Treat (ITT) estimates found almost a 30 percent increase in BA receipt.

**Figure 6. Impact of Upward Bound (UB) on Bachelor's (BA) degree attainment among low-income and first-generation college applicants to Upward Bound: estimates based on 66 of 67 projects in UB sample: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**




*/**/***/**** Significant at 0.10/0.05/.01/00 level.

NOTE: TOT = Treatment on the Treated; ITT= Intent to Treat; EHSGY = Expected High School Graduation Year; NSC = National Student Clearinghouse; SFA = Student Financial Aid. Estimates based on 66 of 67 projects in sample representing 74 percent of UB at the time of the study. One project removed due to introducing bias into estimates in favor of the control group and representational issues. Model based estimates based on STATA logistic and instrumental variables regression taking into account the complex sample design. We use a 2-stage instrumental variables regression procedure to control for selection effects for the Treatment on the Treated (TOT) impact estimates. ITT estimates include 14 percent of control group who were in Upward Bound Math Science or UB and 20-26 percent of treatment group who did not enter Upward Bound. Calculated January 2010.

**Analysis of Control Group Receipt of Alternative Services and Treatment Group Non-Entrance into the Upward Bound Program**

***The majority of the control group also received some form of supplemental pre-college supplemental access services. Most often this was another federal program services such a Talent Search or Upward Bound Math Science***

Before concluding this report another WWC Standard and key issue needs to be discussed. A major standard of the What Works Clearinghouse and of the random assignment method generally is that the treatment and control group must differ on receipt of the intervention or "the treatment" and that the impact must be attributable to the intervention or no conclusion can be reached. From the beginning of the Upward Bound evaluation, concerns have been raised by participating sites that a large percentage of the control group also had pre-college supplemental services, most frequently other Federal TRIO programs such as Talent Search and even in some cases Upward Bound Math Science—a form of Upward Bound itself. They also reported that often those not randomly selected for the UB treatment group were placed in some other similar service as a substitute for not being randomly selected to be given the UB opportunity.

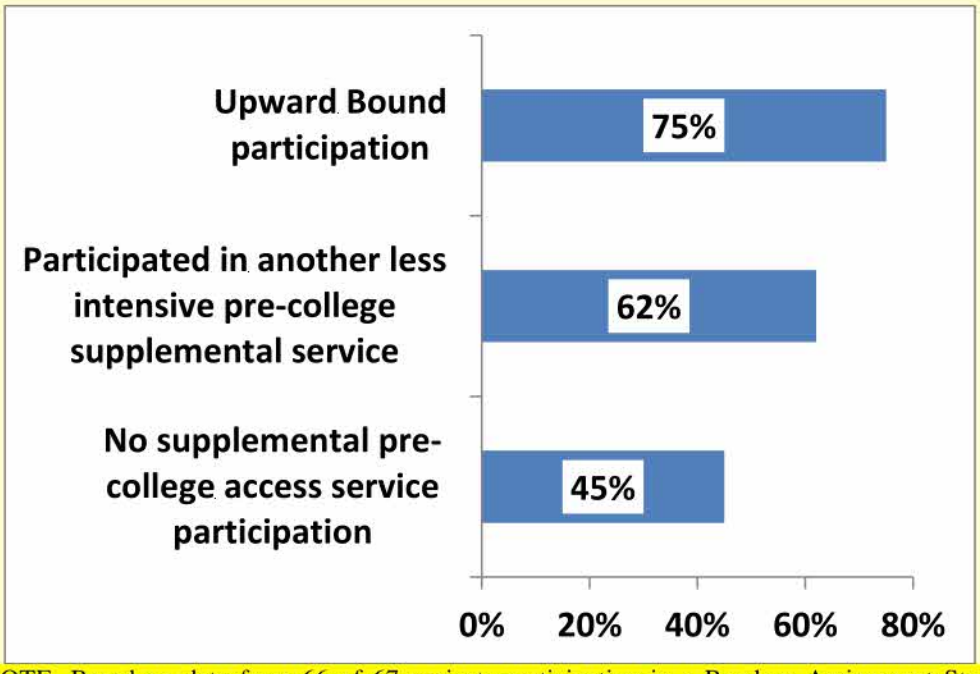
**Extent of Receipt of Pre-College Services among the UB Sample**. An analysis of the random assignment file, baseline and five follow-up surveys reveals key information about the extent to which the sample members from both the treatment and control group participated in various supplemental pre-college services. The random assignment file reveals that about 26 percent the students randomly assigned to be invited into Upward Bound, were coded as "waiting list dropouts." All of these cases were kept in the Intent to Treat (ITT) analyses as Treatment cases although it is unclear as to whether most of these students were actually given the "UB opportunity" due to low-income family mobility and other factors. About 20 percent of the Treatment group reported on the First Follow-up Survey that they never entered Upward Bound and a number could not remember being asked to participate. Although about 20-25 percent of the treatment sample did not enter Upward Bound, overall about 92 percent of the treatment group reported receiving some form of supplemental pre-college services (Upward Bound, Upward Bound Math Science, or some other service such as Talent Search). Conversely among the control group about 14 percent reported entering Upward Bund or Upward Bound Math-Science and overall 60 percent of the control group reported some form of supplemental pre-college services in middle or high school by the end of high school. Most frequently for the control group this was reported to be the less intensive federal service, Talent Search. About one-third of both the treatment and control group reported in study surveys that they received supplemental pre-college services such as Talent Search prior to the Random Assignment.

Surprisingly, even well-known scholars such as Haskins and Rouse (2013) misunderstand the information from the Mathematica study, assuming because of its random assignment method that it is a valid indicator of the effectiveness of all college access programs. This conclusion reflects a lack of understanding of the Upward Bound study and is a misuse of the data. As discussed above, the majority of both the treatment and control group in this study had some form of supplemental pre-college services. As noted in most cases the control group had another federal TRIO service such as Talent Search or Upward Bound Math Science. As noted by

Heckman, Hohman, Smith and Khoo (2000), "evidence that one program is ineffective relative to close substitutes is not evidence that the type of service provided by all of the programs is ineffective, although that is the way experimental evidence is often interpreted." Considered in this light, some of the internal and external reviewers noted that the Mathematica Upward Bound study might be better analyzed using statistical methods such as two stage instrumental variables regression to observe differences in outcome measures for those who participated in different levels of services.

Below we present results observing differences in outcome variables for three groups: 1) those participating in Upward Bound or Upward Bound Math Science; 2) those participating in some other presumably less intensive pre-college (most frequently the federal Talent Search program); and 3) those reporting not receiving any supplemental pre-college services. A two-stage instrumental variables method was used in which the first stage modeled selection differences between these groups on baseline variables and then these factors were used as control variables in the final models. Figures 7 and 8 respectively present results for postsecondary entrance within one year and for award of BA degree in six years for each of the service groups. Similar impacts were also found for financial aid indicators. As seen in Figure 7, about 75 percent of UB participants entered postsecondary education within one year of expected high school graduation. This compares with 45 percent for students reporting no supplemental service college access services participation and 62 percent for those reporting receiving presumably less- intensive supplemental pre-college services.

**Figure 7: Estimates of relative impact of participation in various levels of pre-college access supplemental services on entry into postsecondary education within one year after expected high school graduation: National Evaluation of Upward Bound**



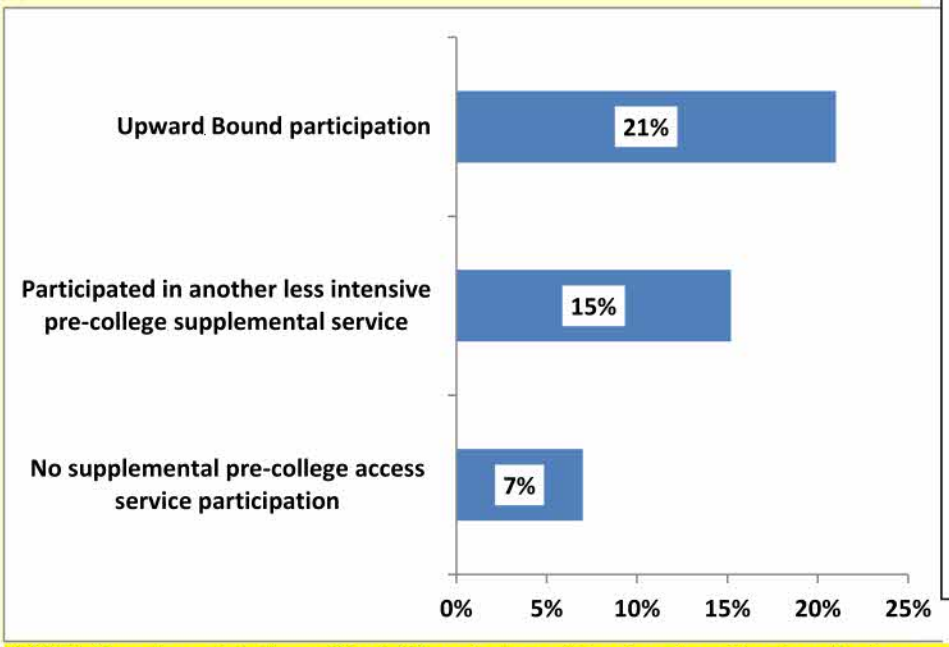

NOTE: Based on data from 66 of 67 projects participating in a Random Assignment Study of about 3,000 middle school and early high school low-income and first-generation UB applicants. The estimates in the figures shown are based on longitudinal data over a 10- year period in an analysis using instrumental two-stage regressions that first model factors related to differences in participation in services and then use these factors in the second stage to control for participation selection bias factors.

SOURCE: Cahalan, Margaret: *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change?* The report can be accessed at the following site: http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml. The study uses National Evaluation of Upward Bound data files and was sponsored by the Policy and Program Studies Services (PPSS) of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education. Study conducted 1992–99 to 2003–04

As Figure 8 below indicates, among those low-income sample members who reported receiving no pre-college supplemental services, about 7 percent were found to have received a BA degree within six years of their expected high school graduation date. This is very similar to the national data from the National Educational Longitudinal Study (NELS) from the same time period (Ingles et.al. 2002) and also Census Bureau data on the percent of students from families in the lowest income quartile who attain a BA by age 24 (about 7 percent in 2004). Among those sample members not receiving Upward Bound or Upward Bound Math Science (UBMS) but reporting receiving some other type of less intensive services such as Talent Search, about 15 percent had achieved a BA degree by six years after their expected high school graduation. Among those who entered the UB or UBMS program, about 21 percent had attained a BA by six years after the expected high school graduation date (Cahalan, 2009). Thus the instrumental variables regression controlling for selection factors revealed that UB participants were 3.3 times

more likely to obtain a BA in six years when compared to those reporting no participation in college access services and 1.4 times as likely when compared to those who reported participating in other presumable less intensive services.

**Figure 8. Estimates of relative impact of participation in various levels of pre-college access supplemental services on BA attainment within 6 years of expected high school graduation: National Evaluation of Upward Bound**




***UB participants were 3.3 times more likely to obtain a BA in six years when compared to those reporting no participation in college access services and 1.4 times as likely when compared to those who reported receiving less intensive services.***

NOTE: Based on data from 66 of 67 projects participating in a Random Assignment Study of about 3,000 middle school and early high school low-income and first-generation UB applicants. The estimates in the figures shown are based on longitudinal data over a 10-year period in an analysis using instrumental two-stage regressions that first model factors related to differences in participation in services and then use these factors in the second stage to control for participation selection bias factors

SOURCE: Cahalan, Margaret: *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change?* The report can be accessed at the following site: http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml. The study uses National Evaluation of Upward Bound data files and was sponsored by the Policy and Program Studies Services (PPSS) of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education. Study conducted 1992–99 to 2003–04.

## Conclusion

Although Mathematica project staff and leadership were sent these fully-documented results in the period of the ED review process of their own final report, and asked to address the concerns raised in the QA review, the results presented above in figures 5 to 8 are not acknowledged in the Mathematica reports. Nor are the seriousness of the representational issues with project 69 or the extent of the treatment control group non-equivalency acknowledged. All impact estimates in the Mathematica reports include project 69, and misleadingly state that the major conclusions do not change substantially because of project 69. Buried in their final report is an admission that results are sensitive to project 69. The report states: "*Because Project 69 had below average impacts, reducing its weight relative to other projects resulted in larger overall impacts for most outcomes compared with the findings from the main impact analysis, which weighted all sample members according to their actual selection probabilities.*". This, however, is also a misleading statement about the effectiveness of project 69. As noted above in Figure 2, a closer look at project 69's treatment and control group clearly reveals that the so-called "below average impacts" in this project were not due to "project 69's poor performance" but were due in fact to the extreme differences between the treatment and control group in favor of the control group in this project.

In summary, as Technical Monitors for the study in QA analyses we found that the Mathematica reports are not transparent in reporting study issues and more robust positive results for Upward Bound. Despite being shown "more credible" positive results for Upward Bound that have been replicated, Mathematica continues to report to Congress, the policy research community, and the public unwarranted and non-transparent conclusions concerning the UB program's effectiveness.[1] This is a very serious matter that needs correcting by Mathematica Policy Research as responsible evaluation contractors and by the US Department of Education.

As noted in 2012, the COE submitted a detailed *Request for Correction* to the US Department of Education. The full text of this request is available at http://www.coenet.us/files/pubs_reports-COE_Request_for_Correction_011712.pdf . As of early 2014, the US Department of Education has refused to consider the COE *Request for Correction* of the Mathematica report, despite the fact that the request was accompanied by an *Statement of Concern* signed by leading researchers that can be found at http://www.coenet.us/files/ED-Statement_of_Concern_011712.pdf. In March of 2014, the co-authors of this paper formally submitted a request to the WWC to rescind its rating of the Mathematica reports as "meets evidence standards without reservations." We now offer this paper in additional support of these two requests.

---

[1] In his Nov 19, 2013 Presidential Address to the Association for Public Policy Analysis and Management (APPAM), Mathematica President and CEO, Dr. Paul Decker, presented the flawed data from the 2009 report (Sefter, et. al. 2009) to reaffirm publically that the UB evaluation study detected no average impacts on UB major legislative goals. He characterized the response of what he called the "Youth Advocacy Community" to the study as constituting "misdemeanors" and "felonies."

**References:**


Cahalan, M. *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change?* can be accessed at the following site http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml

Haskins, R. and Rouse, C. "Time for Change: A New Federal Strategy to Prepare Disadvantaged Students for College," Brookings 2013.

Heckman, J., Hohmann, N, Smith J., and Khoo, M. "Substitution and Dropout Bias in Social Experiments: A Study of an Influential Social Experiment," *The Quarterly Journal of Economics*, May 2000.

Horn, L. J., Chen, X., and MPR Associates. "Toward Resiliency: At-Risk Students Who Make It to College." U.S. Department of Education, Office of Educational Research and Improvement. Washington, DC: U.S. Government Printing Office, May 1998.

*IES, National Center for Education Statistics (NCES) Statistical Standards*--- These may be accessed at the following site url: http://nces.ed.gov/statprog/

*IES, What Works Clearinghouse Standards* ---these may be accessed at the following site urls: http://ies.ed.gov/ncee/wwc/pdf/wwc_version1_standards.pdf
http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=19&tocid=1/

Ingels, S.J., T.R. Curtin, P. Kaufman, M.N. Alt, and Chen, X. *Coming of Age in the1990s: The Eighth-Grade Class of 1988 12 Years Later*. (NCES 2002–321). Washington, DC: U.S. Department of Education, National Center for Education Statistics, 2002.

Myers, D., and Schirm, A. *The Short-Term Impacts of Upward Bound: An Interim Report.* Washington, DC: U.S. Department of Education, Planning and Evaluation Service, 1996.

Myers, D.,, t Olsen, R., Seftor, N., Young, J., and Tuttle, C. "The Impacts of Regular Upward Bound: Results from the Third Follow-Up Data Collection." Report submitted to the U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc., 2004.

Myers, D., and Schirm, A. "The Impacts of Upward Bound: Final Report on Phase I of the National Evaluation." Report submitted to the U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc., 1999.

Nathan, A.B. *Does Upward Bound Have an Effect on Student Educational Outcomes? A Reanalysis of the Horizons Randomized Controlled Trial Study*. A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy (Educational Leadership and Policy Analysis) at the University of Wisconsin-Madison 2013 Date of final oral examination: 02/08/2013.

Olsen, R., Seftor, N., Silva, T., Myers, D, DesRoches, D., and Young, J. "Upward Bound Math Science: Program Description and Interim Impact Estimates." U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc., 2007.

Seftor, Neil S., Arif, M. and Schirm, A.. "*The Impacts of Regular Upward Bound on Postsecondary Outcomes 7-9 Years After Scheduled High School Graduation*." Report submitted to the U.S. Department of Education. Washington, DC: Mathematica Policy Research, Inc., 2009.

Seastrom, M. *NCES Statistical Standards* (NCES 2003–601). U.S. Department of Education, National Center for Education Statistics. Washington, DC: U.S. Government Printing Office, 2002.

*Joint Committee on Standards for Educational Evaluation (JCSEE).* (widely recognized education evaluation professional standards) (website for which is http://www.jcsee.org/)

# Attachment C: Additional Documentation and Statistical Output from Impact Models

**Below are Appendices B and D from the Report *Addressing Study Error in the Random Assignment National Evaluation of Upward Bound: Do the Conclusions Change? By Margaret Cahalan* a COE report published in 2009 and available at http://www.pellinstitute.org/publications-Do_the_Conclusions_Change_2009.shtml.**

## Appendix B

## Examples of Detailed Model Results for Tables in the Body of Report and Comparison of Results when an Alternative Variable is Used for Standardization by EHSGY

Appendix tables B1-B6 provide examples of the complete model results for the statistics reported in the body of the report. We also include some model results (labeled B1a to B4a) for an alternative variable used for establishing a "grade-year" reference for standardization by expected high school graduation year (EHSGY).

As observed in table 2 of the body of the report, treatment and control group non-equivalencies in grades reported on the study surveys indicated that there was a need to standardize outcomes relative to fixed time frames. Model results reported in the body of this paper are based on a standardization of Expected High School Graduation Year (EHSGY) based on a baseline survey variable (B1) present for 99 percent of the sample that read:

> "What grade were you in during the LAST SCHOOL YEAR (1992-93 school year)?"

Because some students reportedly answered the question with reference to the 1991-92 school year instead of 1992-93, we included a correction in the tabulation that provided for a range going from -1 to +1 (or +4) of the year that was established on the basis of the grade reported on the baseline data file which ranged from grade 7 to a few in 11. As discussed in the body of the report and indicated in tables 2 and 3, estimates of EHSGY using different variables on the data files are not entirely consistent with each other when the baseline survey, first follow-up survey and the third follow-up survey data files are compared. For this reason, to check the models reported in the body of the text based on the baseline variable (B1) with the correction for the 1991-92 responders, we also calculated an alternate EHSGY based on the results to the first follow-up survey in which sample members were asked the following question (QA1):

> "What grade (are you in/were you in during the 1993-94 school year) or (are/were) you not attending junior high or high school (now/then)?"

The models presented in appendix tables B1 to B4a show comparative results using the two alternative variables for tabulation of EHSGY. As can be observed in the tables, the two alternative bases for standardization yield much the same impact estimates and significance test results. For example, table B1 and B1a present the same model with the two alternative variables used to standardize for the outcome of postsecondary entrance evidence by +1 of EHSGY.

The impact estimate reported in table B-1 shows a 6.9 effect significant at the .004 level. These estimates are used in the body of the report. The alternative variable presented in table B1a shows a 6.7 effect significant at the .000 level for the same model. Results in tables B2 and B2a are for the same models as in B1 and B1a but exclude project 69. We see that the estimate of effect in B2 (and reported in table 6 in the body of the report) is 9.1 significant at the .000 level and the estimate in B2a using the alternative first follow-up variable for standardization shows an effect size of 12.7 and significance of .001. Tables B3 and B3a show instrumental variables two stage regression results for modeling TOT with the dependent variable of appearance on the aid file and show effect size of 9.3 significant at the .002 level for the baseline variable standardization; and show an effect size of 10.4 significant at the .001 level for the first follow- up variable used for standardization. Results in B4 and B4a for bachelor's degree receipt without project 69 show similar effects and significance levels when +7 instead of +6 is used with the baseline.

**Table B-1. Intent to Treat (ITT) logistic regression results models for dependent variable of having evidence of postsecondary from any applicable survey or from SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (estimate reported in table 5 and 6)**

| pr-T = 72.9<br>pr-C = 66.0<br>Difference = 6.9****<br>kenye2 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | Ffutc | 0.395308 | 0.130178 | 3.04 | 0.004 | 0.131997 | 0.658619 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 0.165761 | 0.668542 | 0.25 | 0.805 | -1.18649 | 1.518015 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.80596 | 0.446533 | -1.8 | 0.079 | -1.70916 | 0.097236 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 0.223571 | 0.848204 | 0.26 | 0.793 | -1.49208 | 1.939225 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -1.2639 | 1.263728 | -1 | 0.323 | -3.82003 | 1.292227 |
| Clowoy (Low income only) | Clowoy | 0.189757 | 0.257292 | 0.74 | 0.465 | -0.33066 | 0.710179 |
| Cfgenoy (First generation only) | Cfgenoy | 0.346913 | 0.212268 | 1.63 | 0.11 | -0.08244 | 0.776266 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.96561 | 1.159979 | -0.83 | 0.41 | -3.31189 | 1.380672 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.36939 | 0.391458 | -0.94 | 0.351 | -1.16119 | 0.422409 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.117 | 0.638178 | -1.75 | 0.088 | -2.40784 | 0.173837 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.7174 | 0.136398 | -5.26 | 0 | -0.99329 | -0.44151 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -1.15535 | 0.263932 | -4.38 | 0 | -1.68921 | -0.6215 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -1.08164 | 0.159795 | -6.77 | 0 | -1.40485 | -0.75842 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.62101 | 0.096702 | -6.42 | 0 | -0.81661 | -0.42541 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.130731 | 0.127043 | 1.03 | 0.31 | -0.12624 | 0.387699 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | 0.260035 | 0.125456 | 2.07 | 0.045 | 0.006277 | 0.513794 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | Cothrac | -0.11733 | 0.298544 | -0.39 | 0.696 | -0.72119 | 0.486535 |
| Chisp (Hispanic—ref Black) | Chisp | -0.3342 | 0.21233 | -1.57 | 0.124 | -0.76368 | 0.095275 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.50434 | 0.164489 | -3.07 | 0.004 | -0.83705 | -0.17163 |
| Cfemale (Female) | Cfemale | 0.655618 | 0.074893 | 8.75 | 0 | 0.504132 | 0.807103 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.404186 | 0.15019 | 2.69 | 0.01 | 0.100399 | 0.707974 |
| _cons | _cons | 0.983643 | 0.520775 | 1.89 | 0.066 | -0.06972 | 2.03701 |

NOTE: Results of this table appear in figure I and in table 5 and table 6. Standardized based on baseline survey question B1 with correction for 1991-92 responders. SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 for additional note information. See also table B-1a for results using an alternative variable for EHSGY estimation. Number of strata (wprstco)= 28; Number of PSU ( wprojid) = 67; uses poststratified longitudinal baseline weight (v5bwgtp1).

**SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-1a. Intent to Treat (ITT) logistic regression results models for dependent variable of having evidence of postsecondary from any applicable survey or from SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (Uses alternative grade variable for standardization of EHSGY)**

| pr-T = 72.7<br>pr-C = 66.0<br>Difference = 6.7****.<br>npse18 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.371381 | 0.092769 | 4 | 0 | 0.183739 | 0.559023 |
| Ffgr9 (grade 10 ref) | ffgr9 | 0.00818 | 0.329167 | 0.02 | 0.98 | -0.65762 | 0.673983 |
| Ffgr11 (grade 10 ref) | ffgr11 | -0.16084 | 0.164541 | -0.98 | 0.334 | -0.49366 | 0.17197 |
| Ffgr12 (grade 10 ref) | ffgr12 | -1.62816 | 0.395105 | -4.12 | 0 | -2.42734 | -0.82898 |
| Clowoy (Low income only) | clowoy | 0.274296 | 0.247483 | 1.11 | 0.275 | -0.22629 | 0.774878 |
| Cfgenoy (First generation only) | cfgenoy | 0.375038 | 0.193827 | 1.93 | 0.06 | -0.01701 | 0.76709 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.783521 | 0.301535 | 2.6 | 0.013 | 0.17361 | 1.393432 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.427427 | 0.226795 | 1.88 | 0.067 | -0.03131 | 0.886164 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.31887 | 0.207559 | -1.54 | 0.133 | -0.7387 | 0.100957 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.79806 | 0.145596 | -5.48 | 0 | -1.09256 | -0.50357 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.28867 | 0.286644 | -4.5 | 0 | -1.86846 | -0.70888 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.96014 | 0.163404 | -5.88 | 0 | -1.29065 | -0.62962 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.62401 | 0.119286 | -5.23 | 0 | -0.86529 | -0.38273 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.035613 | 0.11215 | 0.32 | 0.753 | -0.19123 | 0.262458 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.223459 | 0.101548 | 2.2 | 0.034 | 0.01806 | 0.428859 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.0687 | 0.285566 | -0.24 | 0.811 | -0.64631 | 0.50891 |
| Chisp (Hispanic—ref Black) | chisp | -0.26647 | 0.196386 | -1.36 | 0.183 | -0.6637 | 0.130757 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.5774 | 0.210339 | -2.75 | 0.009 | -1.00285 | -0.15195 |
| Cfemale (Female) | cfemale | 0.578733 | 0.077784 | 7.44 | 0 | 0.421399 | 0.736067 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.400858 | 0.131542 | 3.05 | 0.004 | 0.134789 | 0.666926 |
| _cons | _cons | | | | | | |

NOTE: Model uses an alternative variable from the First Follow-up (A3) instead of variable B1 from the Baseline Survey on which to standardize EHSGY (See table B-1). SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 in body of text for additional note information. Number of strata (wprstco)= 28; Number of PSU ( wprojid) = 67; uses postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE:** Data tabulated June 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-2. Intent to Treat (ITT), *excludes project 69,* logistic regression results models for dependent variable of having evidence of postsecondary from any applicable survey or from SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (estimate reported in table 5 and table 6)**

| pr-T = 73.4<br>pr-C = 64.3<br>Difference = 9.1****<br>kenye2 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.489536 | 0.128311 | 3.82 | 0 | 0.230002 | 0.74907 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 0.521967 | 0.625101 | 0.84 | 0.409 | -0.74242 | 1.786352 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.39121 | 0.286428 | -1.37 | 0.18 | -0.97057 | 0.188144 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.45553 | 0.563581 | -0.81 | 0.424 | -1.59548 | 0.68442 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -2.12228 | 0.983931 | -2.16 | 0.037 | -4.11247 | -0.1321 |
| Clowoy (Low income only) | clowoy | 0.346651 | 0.249567 | 1.39 | 0.173 | -0.15815 | 0.851449 |
| Cfgenoy (First generation only) | cfgenoy | 0.456902 | 0.247804 | 1.84 | 0.073 | -0.04433 | 0.958132 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.152517 | 0.657632 | 0.23 | 0.818 | -1.17767 | 1.482704 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.10336 | 0.320505 | -0.32 | 0.749 | -0.75164 | 0.544922 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.08313 | 0.627791 | -1.73 | 0.092 | -2.35296 | 0.186695 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.79163 | 0.149155 | -5.31 | 0 | -1.09333 | -0.48994 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.24988 | 0.364853 | -3.43 | 0.001 | -1.98787 | -0.5119 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.96619 | 0.175638 | -5.5 | 0 | -1.32145 | -0.61092 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.65279 | 0.123614 | -5.28 | 0 | -0.90282 | -0.40276 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.162114 | 0.149791 | 1.08 | 0.286 | -0.14087 | 0.465095 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.346946 | 0.121938 | 2.85 | 0.007 | 0.100304 | 0.593588 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.062038 | 0.279237 | 0.22 | 0.825 | -0.50277 | 0.626848 |
| Chisp (Hispanic—ref Black) | chisp | -0.28652 | 0.3381 | -0.85 | 0.402 | -0.97039 | 0.397353 |
| Cwhite (Rac e was White, not Hispanic—ref Black) | cwhite | -0.45507 | 0.167869 | -2.71 | 0.01 | -0.79462 | -0.11552 |
| Cfemale (Female) | cfemale | 0.651833 | 0.088637 | 7.35 | 0 | 0.472549 | 0.831118 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.288657 | 0.175721 | 1.64 | 0.108 | -0.06677 | 0.644085 |
| _cons | _cons | 0.519335 | 0.33867 | 1.53 | 0.133 | -0.16569 | 1.20436 |

NOTE: Results of this table appear in figure I and in table 5 and table 6. Standardized based on Baseline Survey question B1 with correction for 1991-92 responders. SFA = Student Financial Aid files; Ref = left out reference in dummy variable sequence. See table 5 for additional note information. See also table B-2a for results using an alternative variable for EHSGY estimation. Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE:** Data tabulated June 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-2a. Intent to Treat (ITT), *excludes project 69,* logistic regression results models for dependent variable of having evidence of postsecondary from any applicable survey or from SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (uses alternative grade variable for EHSGY standardization)**

| pr-T = 73.3<br>pr-C = 60.6<br>Difference = 12.7****<br>npse18 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.404011 | 0.113939 | 3.55 | 0.001 | 0.173547 | 0.634475 |
| Ffgr9 (grade 10 ref) | ffgr9 | 0.37182 | 0.280346 | 1.33 | 0.192 | -0.19523 | 0.938873 |
| Ffgr11 (grade 10 ref) | ffgr11 | -0.02695 | 0.159631 | -0.17 | 0.867 | -0.34983 | 0.295939 |
| Ffgr12 (grade 10 ref) | ffgr12 | -1.32323 | 0.547936 | -2.41 | 0.021 | -2.43154 | -0.21493 |
| Clowoy (Low income only) | clowoy | 0.465974 | 0.202402 | 2.3 | 0.027 | 0.056577 | 0.87537 |
| Cfgenoy (First generation only) | cfgenoy | 0.46034 | 0.253056 | 1.82 | 0.077 | -0.05151 | 0.972195 |
| C11gssf . (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.663936 | 0.314424 | 2.11 | 0.041 | 0.027953 | 1.299919 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.358413 | 0.261897 | 1.37 | 0.179 | -0.17132 | 0.888151 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.4937 | 0.208163 | -2.37 | 0.023 | -0.91475 | -0.07265 |
| Cexdk (Baseline educational expectation was don't know---ref BA) | cexdk | -0.853 | 0.158962 | -5.37 | 0 | -1.17454 | -0.53147 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.38008 | 0.38735 | -3.56 | 0.001 | -2.16357 | -0.59659 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.89794 | 0.182897 | -4.91 | 0 | -1.26788 | -0.528 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.67054 | 0.141661 | -4.73 | 0 | -0.95708 | -0.38401 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA). | cexma | 0.028171 | 0.146501 | 0.19 | 0.849 | -0.26816 | 0.324498 |
| Cexphd (Baseline educational expectation was PhD---ref BA). | cexphd | 0.28854 | 0.102476 | 2.82 | 0.008 | 0.081263 | 0.495816 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black). | cothrac | 0.189738 | 0.214378 | 0.89 | 0.382 | -0.24388 | 0.623358 |
| Chisp (Hispanic—ref Black) | chisp | -0.20354 | 0.293326 | -0.69 | 0.492 | -0.79685 | 0.389766 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.44839 | 0.182601 | -2.46 | 0.019 | -0.81774 | -0.07905 |
| Cfemale (Female) | cfemale | 0.583567 | 0.095933 | 6.08 | 0 | 0.389524 | 0.77761 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.273002 | 0.128501 | 2.12 | 0.04 | 0.013084 | 0.532919 |
| _cons | _cons | 0.099075 | 0.269472 | 0.37 | 0.715 | -0.44598 | 0.644134 |

NOTE: Model uses an alternative variable from the First Follow-up (A3) instead of variable B1 from the Baseline Survey on which to standardize EHSGY (See table B-2). SFA = Student Financial Aid files; Ref = left out reference in dummy variable sequence. See table 5 in text for additional note information; Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE:** Data tabulated June 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-3. Instrumental variables regression for Treated on Treated (TOT) modeling dependent variable of appearing on the federal SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (estimate reported in table 8)**

| xb T = 62.9<br>xb C = 53.6<br>Difference = 9.3****<br>(Kaidhs) | Variable name | Coef. | Linearized Std. Err. | T | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| xnewgp (Participated in UB/UBMS) | xnewgp | 0.123149 | 0.037028 | 3.33 | 0.002 | 0.048253 | 0.198045 |
| Gr79293 (Grade 7 in 1992-93 ref grade 9) | gr79293 | -0.01293 | 0.132541 | -0.1 | 0.923 | -0.28102 | 0.255162 |
| Gr89293 (Grade 8 in 1992-93 ref grade 9) | gr89293 | -0.12538 | 0.048361 | -2.59 | 0.013 | -0.2232 | -0.02756 |
| Gr109293 (Grade 10 in 1992-93 ref grade 9) | gr109293 | 0.008279 | 0.125262 | 0.07 | 0.948 | -0.24509 | 0.261646 |
| Gr119293 (Grade 11 in 1992-93 ref grade 9) | gr119293 | -0.24429 | 0.168436 | -1.45 | 0.155 | -0.58498 | 0.096407 |
| Clowoy (Low income only) | clowoy | 0.017819 | 0.049742 | 0.36 | 0.722 | -0.08279 | 0.118432 |
| Cfgenoy (First generation only) | cfgenoy | 0.051589 | 0.051158 | 1.01 | 0.319 | -0.05189 | 0.155066 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.02485 | 0.141647 | -0.18 | 0.862 | -0.31135 | 0.261663 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.04074 | 0.055938 | -0.73 | 0.471 | -0.15389 | 0.072402 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.14285 | 0.118341 | -1.21 | 0.235 | -0.38222 | 0.096515 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.18389 | 0.040305 | -4.56 | 0 | -0.26541 | -0.10236 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -0.26295 | 0.0628 | -4.19 | 0 | -0.38997 | -0.13592 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.18834 | 0.035934 | -5.24 | 0 | -0.26103 | -0.11566 |
| cex14aa (Baseline educational expectation was 2-year---ref BA) | cex14aa | -0.17929 | 0.024134 | -7.43 | 0 | -0.22811 | -0.13047 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.022336 | 0.025866 | 0.86 | 0.393 | -0.02998 | 0.074655 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.042772 | 0.018322 | 2.33 | 0.025 | 0.005713 | 0.079831 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.018877 | 0.046879 | 0.4 | 0.689 | -0.07595 | 0.113699 |
| Chisp (Hispanic—ref Black) | chisp | -0.06829 | 0.06105 | -1.12 | 0.27 | -0.19177 | 0.055199 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.10059 | 0.038615 | -2.6 | 0.013 | -0.17869 | -0.02248 |
| Cfemale (Female) | cfemale | 0.139323 | 0.02847 | 4.89 | 0 | 0.081736 | 0.196909 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.032356 | 0.0231 | 1.4 | 0.169 | -0.01437 | 0.07908 |
| _cons | _cons | 0.586899 | 0.075209 | 7.8 | 0 | 0.434774 | 0.739023 |

NOTE: Results of this table appear in table 8. Standardized based on baseline survey question B1 with correction for 1991-92 responders. SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 for additional note information. See also table B-3a for results using an alternative variable for EHSGY estimation. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; uses poststratified longitudinal baseline weight (v5bwgtp1). Instrumented: xnewgp; Instruments: gr79293 gr89293 gr109293 gr119293 clowoy cfgenoy c11gssf c10gssf c8gssfm cexdk cex13v cexhs cex14aa cexma cexphd cothrac chisp cwhite cfemale parbefor ffutc.

**SOURCE:** Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-3a. Instrumental variables regression for Treated on Treated (TOT) modeling dependent variable of appearing on the federal SFA files by +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (uses alternative grade variable for standardization)**

| pr-T = 65.4<br>pr-C = 55.0<br>Difference = 10.4****<br>knaidh1 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| xnewgp (Participated in UB/UBMS) | xnewgp | 0.098531 | 0.02724 | 3.62 | 0.001 | 0.043433 | 0.153629 |
| Ffgr9 (grade 10 ref) | ffgr9 | -0.00902 | 0.052581 | -0.17 | 0.865 | -0.11538 | 0.097332 |
| Ffgr11 (grade 10 ref) | ffgr11 | 0.014367 | 0.027164 | 0.53 | 0.6 | -0.04058 | 0.069311 |
| Ffgr12 (grade 10 ref) | ffgr12 | -0.24887 | 0.064124 | -3.88 | 0 | -0.37858 | -0.11917 |
| Clowoy (Low income only) | clowoy | 0.057143 | 0.048196 | 1.19 | 0.243 | -0.04034 | 0.154628 |
| Cfgenoy (First generation only) | cfgenoy | 0.029202 | 0.039698 | 0.74 | 0.466 | -0.0511 | 0.109499 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.178713 | 0.056336 | 3.17 | 0.003 | 0.064763 | 0.292663 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.073839 | 0.037345 | 1.98 | 0.055 | -0.0017 | 0.149376 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.01392 | 0.062567 | -0.22 | 0.825 | -0.14047 | 0.112637 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.17051 | 0.031567 | -5.4 | 0 | -0.23436 | -0.10666 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -0.24205 | 0.060803 | -3.98 | 0 | -0.36504 | -0.11907 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.1457 | 0.02996 | -4.86 | 0 | -0.2063 | -0.0851 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.17534 | 0.029116 | -6.02 | 0 | -0.23423 | -0.11645 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.012577 | 0.023624 | 0.53 | 0.597 | -0.03521 | 0.06036 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.014907 | 0.026489 | 0.56 | 0.577 | -0.03867 | 0.068485 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.038572 | 0.0459 | 0.84 | 0.406 | -0.05427 | 0.131412 |
| Chisp (Hispanic—ref Black) | chisp | -0.02733 | 0.036807 | -0.74 | 0.462 | -0.10178 | 0.047119 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.08829 | 0.03744 | -2.36 | 0.023 | -0.16402 | -0.01256 |
| Cfemale (Female) | cfemale | 0.146096 | 0.026613 | 5.49 | 0 | 0.092267 | 0.199925 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.064862 | 0.021253 | 3.05 | 0.004 | 0.021875 | 0.10785 |
| _cons | _cons | 0.457948 | 0.040832 | 11.22 | 0 | 0.375358 | 0.540537 |

NOTE: Model uses an alternative variable from the first follow-up survey (A3) instead of variable B1 from the Baseline Survey on which to standardize EHSGY (See table B-3). SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 in body of text for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; uses poststratified longitudinal baseline weight (v5bwgtp1). Instrumented: xnewgp; Instruments: gr79293 gr89293 gr109293 gr119293 clowoy cfgenoy c11gssf c10gssf c8gssfm cexdk cex13v cexhs cex14aa cexma cexphd cothrac chisp cwhite cfemale parbefor ffutc.

**SOURCE:** Data tabulated June 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; and federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-4 Intent to Treat (ITT), *excludes project 69*, logistic regression results models for dependent variable of having evidence of attaining a BA degree in +6 of expected high school graduation year (EHSGY) from any applicable survey, SFA Files, or National Student Clearinghouse (NSC): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (estimate in table 10)**

| pr-T = 17.0<br>pr-C = 13.3<br>Difference = 3.7****<br>kbahs6 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.362466 | 0.084166 | 4.31 | 0 | 0.192224 | 0.532708 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 0.092829 | 0.334572 | 0.28 | 0.783 | -0.58391 | 0.769564 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr99293 | 0.446231 | 0.304951 | 1.46 | 0.151 | -0.17059 | 1.063053 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 0.474865 | 0.600971 | 0.79 | 0.434 | -0.74071 | 1.690443 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -1.46837 | 1.670039 | -0.88 | 0.385 | -4.84634 | 1.909606 |
| Clowoy . (Low income only) | clowoy | 0.706581 | 0.265714 | 2.66 | 0.011 | 0.169123 | 1.244038 |
| Cfgenoy (First generation only) | cfgenoy | 0.534226 | 0.191623 | 2.79 | 0.008 | 0.146633 | 0.92182 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.40178 | 0.652044 | -0.62 | 0.541 | -1.72066 | 0.917103 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.38824 | 0.265334 | -1.46 | 0.151 | -0.92493 | 0.148449 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.53442 | 0.4097 | -1.3 | 0.2 | -1.36312 | 0.294275 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.67588 | 0.215566 | -3.14 | 0.003 | -1.1119 | -0.23985 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -2.17255 | 0.908448 | -2.39 | 0.022 | -4.01006 | -0.33504 |
| Cex13v. (Baseline educational expectation was vocational---ref BA) | cex13v | -0.6227 | 0.277369 | -2.25 | 0.031 | -1.18374 | -0.06167 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -1.28374 | 0.274614 | -4.67 | 0 | -1.8392 | -0.72828 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.250644 | 0.165068 | 1.52 | 0.137 | -0.08324 | 0.584526 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.19915 | 0.158047 | 1.26 | 0.215 | -0.12053 | 0.518831 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.421884 | 0.268605 | 1.57 | 0.124 | -0.12142 | 0.965189 |
| Chisp (Hispanic—ref Black) | chisp | -0.15843 | 0.244249 | -0.65 | 0.52 | -0.65247 | 0.335611 |
| Cwhite (Race was White, not Hispanic—ref Black). | cwhite | -0.25651 | 0.169433 | -1.51 | 0.138 | -0.59922 | 0.086198 |
| Cfemale (Female) | cfemale | 0.662424 | 0.125721 | 5.27 | 0 | 0.40813 | 0.916719 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.116322 | 0.110851 | 1.05 | 0.3 | -0.10789 | 0.340539 |
| _cons | _cons | -2.22556 | 0.185144 | -12.02 | 0 | -2.60005 | -1.85107 |

NOTE: Results of this table appear in table 10. Standardized based on baseline survey question B1 with correction for 1991-92 responders. SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 for additional note information. See also table B-4a for results using an alternative variable for EHSGY estimation. Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE**:. Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04; and National Student Clearinghouse Data 1995-2004

**Table B-4a Intent to Treat (ITT), *excludes project 69*, logistic regression results models for dependent variable of having evidence of attaining a BA degree in +7 of expected high school graduation year (EHSGY) from the any applicable survey, SFA Files, or National Student Clearinghouse (NSC): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04 (uses alternative grade variable for standardization)**

| pr-T = 18.0<br>pr-C = 14.5<br>Difference = 3.5****<br>knba7 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.330297 | 0.082796 | 3.99 | 0 | 0.162826 | 0.497767 |
| Ffgr9 (grade 10 ref) | ffgr9 | -0.22032 | 0.215644 | -1.02 | 0.313 | -0.6565 | 0.215859 |
| Ffgr11 (grade 10 ref) | ffgr11 | -0.00821 | 0.246862 | -0.03 | 0.974 | -0.50754 | 0.491114 |
| Ffgr12 (grade 10 ref) | ffgr12 | -1.20172 | 0.617098 | -1.95 | 0.059 | -2.44992 | 0.046477 |
| Clowoy (Low income only) | clowoy | 0.684771 | 0.264543 | 2.59 | 0.013 | 0.149683 | 1.219859 |
| Cfgenoy (First generation only) | cfgenoy | 0.577483 | 0.155877 | 3.7 | 0.001 | 0.262192 | 0.892773 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.341965 | 0.384344 | 0.89 | 0.379 | -0.43544 | 1.119374 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.040414 | 0.169063 | 0.24 | 0.812 | -0.30155 | 0.382376 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.28152 | 0.224888 | -1.25 | 0.218 | -0.7364 | 0.173358 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.59992 | 0.208694 | -2.87 | 0.007 | -1.02204 | -0.1778 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -2.19774 | 0.883071 | -2.49 | 0.017 | -3.98391 | -0.41156 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.53217 | 0.287618 | -1.85 | 0.072 | -1.11393 | 0.049596 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -1.26668 | 0.275738 | -4.59 | 0 | -1.82441 | -0.70895 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.35107 | 0.197926 | 1.77 | 0.084 | -0.04927 | 0.751413 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.213193 | 0.157779 | 1.35 | 0.184 | -0.10595 | 0.532331 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.382197 | 0.25811 | 1.48 | 0.147 | -0.13988 | 0.904273 |
| Chisp (Hispanic—ref Black) | chisp | -0.12792 | 0.225029 | -0.57 | 0.573 | -0.58309 | 0.327239 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.31609 | 0.154333 | -2.05 | 0.047 | -0.62826 | -0.00392 |
| Cfemale (Female) | cfemale | 0.67513 | 0.117384 | 5.75 | 0 | 0.437698 | 0.912561 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.198067 | 0.097729 | 2.03 | 0.05 | 0.000391 | 0.395742 |
| _cons | _cons | -2.10691 | 0.195875 | -10.76 | 0 | -2.50311 | -1.71072 |

NOTE: Model uses an alternative variable from the first follow up (A3) on which to standardize grade. Model uses an alternative variable from the First Follow-up Survey (A3) instead of variable B1 from the Baseline Survey on which to standardize EHSGY (See table B-4). SFA = Student Financial Aid files; Ref = left out reference in dummy variable sequence. See table 5 in body of report for additional note information; Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04; and National Student Clearinghouse Data 1995-2004.

**Table B-5. ITT Logistic Regression model results for higher academic risk subgroup (bottom 20 percent on 9th grade academic indicators) modeling dependent variable of having evidence of entering postsecondary within +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-2004 (estimate in table 12 in text)**

| Pr---Treatment- 60.1<br>. Control- 41.0<br>. Difference-. 19.1***. n3ar20h== 1)<br>521. cases-- kenye2 | Variable name | Coef. | Linearized Std. Err. | T | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| Ffutc (random assignment to treatment) | ffutc | 0.68927 | 0.243843 | 2.83 | 0.007 | 0.196051 | 1.182489 |
| Gr79293 (Grade 7 in 1992-93 ref grade 9) | gr79293 | 2.221241 | 1.032169 | 2.15 | 0.038 | 0.133482 | 4.308999 |
| Gr89293 (Grade 8 in 1992-93 ref grade 9) | gr89293 | -0.25723 | 0.640009 | -0.4 | 0.69 | -1.55177 | 1.037314 |
| Gr109293 (Grade 10 in 1992-93 ref grade 9) | gr109293 | 0.870869 | 0.988209 | 0.88 | 0.384 | -1.12797 | 2.869711 |
| Gr119293 (Grade 11 in 1992-93 ref grade 9) | | | | | | | |
| Clowoy . (Low income only) | clowoy | 0.677017 | 0.435089 | 1.56 | 0.128 | -0.20303 | 1.557068 |
| Cfgenoy (First generation only) | cfgenoy | 0.543434 | 0.497777 | 1.09 | 0.282 | -0.46342 | 1.550283 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -1.34637 | 1.380078 | -0.98 | 0.335 | -4.13784 | 1.445102 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.06583 | 0.682602 | 0.1 | 0.924 | -1.31486 | 1.446522 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -2.55885 | 1.08985 | -2.35 | 0.024 | -4.76328 | -0.35442 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -0.65612 | 0.522212 | -1.26 | 0.216 | -1.71239 | 0.400155 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -2.14998 | 1.264569 | -1.7 | 0.097 | -4.70782 | 0.407849 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.98 | 0.453137 | -2.16 | 0.037 | -1.89656 | -0.06345 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.50108 | 0.412805 | -1.21 | 0.232 | -1.33606 | 0.333896 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA). | cexma | -0.11761 | 0.344832 | -0.34 | 0.735 | -0.8151 | 0.579874 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | -0.80588 | 0.469222 | -1.72 | 0.094 | -1.75497 | 0.143209 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.54198 | 0.572089 | -0.95 | 0.349 | -1.69914 | 0.615182 |
| Chisp (Hispanic—ref Black) | chisp | -0.4907 | 0.643455 | -0.76 | 0.45 | -1.79221 | 0.810811 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.647 | 0.348201 | -1.86 | 0.071 | -1.3513 | 0.057307 |
| Cfemale (Female) | cfemale | 0.534927 | 0.181975 | 2.94 | 0.005 | 0.166847 | 0.903007 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.411237 | 0.334681 | 1.23 | 0.227 | -0.26572 | 1.088194 |
| _cons | _cons | 0.080847 | 0.53191 | 0.15 | 0.88 | -0.99504 | 1.156736 |

NOTE: Results of this table appear in figure 9 and table 12. Standardized based on baseline survey question B1 with correction for 1991-92 responders. SFA = Student Financial Aid files. Ref = left out reference in dummy variable sequence. See table 5 for additional note information. Number of strata (wprstco)= 28; Number of PSU (wprojid) = 67; uses postratified longitudinal baseline weight (v5bwgtp1).

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table B-6. Intent to Treat (ITT) logistic regression results for sample members with lower educational expectations for modeling of dependent variables of attainment of any postsecondary credential using survey data only adjusted for non-response: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| Pr--- Treatment- 50.3<br>. Control-. . 35.0<br>Difference -. . 15.3**<br>bahexp == 0 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | [95% Conf. | Interval] |
|---|---|---|---|---|---|---|---|
| Ffutc (random assignment to treatment) | ffutc | 0.781074 | 0.356626 | 2.19 | 0.035 | 0.059731 | 1.502417 |
| Gr79293 (Grade 7 in 1992-93 ref grade 9) | gr79293 | 15.74681 | 0.737344 | 21.36 | 0 | 14.25539 | 17.23823 |
| Gr89293 (Grade 8 in 1992-93 ref grade 9) | gr89293 | -1.76353 | 0.856123 | -2.06 | 0.046 | -3.4952 | -0.03186 |
| Gr109293 (Grade 10 in 1992-93 ref grade 9) | gr109293 | -0.04512 | 0.81914 | -0.06 | 0.956 | -1.70199 | 1.611743 |
| Gr119293 (Grade 11 in 1992-93 ref grade 9) | | | | | | | |
| Clowoy . (Low income only) | clowoy | 1.004854 | 0.97003 | 1.04 | 0.307 | -0.95722 | 2.966925 |
| Cfgenoy (First generation only) | cfgenoy | 0.316873 | 0.610377 | 0.52 | 0.607 | -0.91773 | 1.551476 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -1.56826 | 1.093469 | -1.43 | 0.159 | -3.78 | 0.643495 |
| C10gssf . (Grade was 10 on student selection form—ref grade 9) | c10gssf | -1.38135 | 0.637629 | -2.17 | 0.036 | -2.67107 | -0.09162 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -17.8409 | . | . | . | . | . |
| Cexdk (Baseline educational expectation was don't know---ref BA) | cexdk | -0.45648 | 0.571858 | -0.8 | 0.43 | -1.61317 | 0.700211 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | 1.464931 | 0.468557 | 3.13 | 0.003 | 0.517185 | 2.412677 |
| Cex13v. (Baseline educational expectation was vocational---ref BA) | | | | | | | |
| cex14aa (Baseline educational expectation was 2-year---ref BA) | cex14aa | -0.71024 | 0.872378 | -0.81 | 0.421 | -2.47479 | 1.054315 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | | | | | | | |
| Cexphd (Baseline educational expectation was Ph.D---ref BA) | | | | | | | |
| Cothrac (Race was not Hispanic or Black or White—ref Black) | cothrac | 0.188601 | 0.435626 | 0.43 | 0.667 | -0.69254 | 1.069738 |
| Chisp (Hispanic—ref Black) | chisp | 0.65849 | 0.522057 | 1.26 | 0.215 | -0.39747 | 1.714449 |
| Cwhite (Race was White, not Hispanic—ref Black). | cwhite | -0.35067 | 0.413444 | -0.85 | 0.402 | -1.18694 | 0.485598 |
| Cfemale (Female) | cfemale | 0.769137 | 0.544524 | 1.41 | 0.166 | -0.33227 | 1.87054 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.135167 | 0.207005 | 0.65 | 0.518 | -0.28354 | 0.553873 |
| _cons | _cons | 0.604448 | 0.873176 | 0.69 | 0.493 | -1.16172 | 2.370613 |

Note: Ref = left out reference in dummy variable sequence. See table 5 in text for complete note information; Number of strata = 28; Number of PSU = 67; Note results using survey data only subject to non-response bias and sub-group results subject to unequal weighting.

# Appendix D
# Additional Tabulations: Sensitivity to Project 69; Standardization and Use of Administrative Records

The tables in the body of the report attempt to mitigate some of the observed bias introduced by project 69 by use of a longitudinal file composed of all sample members rather than only responders to the survey, standardization of outcomes to EHSGY, and care in the use of the National Student Clearinghouse (NSC) data. This appendix addresses the issue of the sensitivity of results to project 69 when these procedures are not followed. Specifically it addresses the question of the sensitivity of the results published by ED in the third follow-up report to inclusion of 69 and also to standardization and use of administrative records. We also include some tables comparing fourth follow-up data with and without 69. We also include a table that compares results from the main statistic on postsecondary entrance from the fifth follow-up report with and without project 69. Note that none of the models included control for the academic differences in treatment and control group introduced by 69, as these were measures from the 9th grade after treatment had begun for a portion of the sample.

Tables D-1 and D-1a present models based on only those who responded to the third follow-up report for the outcome of "any postsecondary enrollment." The weights were adjusted for non-response by project and weighted up to the poststratified totals. The outcome variable is not standardized as to expected high school graduation year (EHSGY) and the file uses only survey data unsupplemented by SFA data. As can be seen, results in D-1 without the outlier project 69 are significant, and those with 69 are not significant (D-1a). Results in D-1a are consistent with those that formed the basis of the conclusions in the published third follow-up report which stated that the program had no effect on postsecondary enrollment.

Tables D-2 and D-2a present similar models from the fourth follow-up survey. The tables are also based on survey responders only, do not include standardization for EHSGY, and do not use SFA records. The models use the fourth follow-up non-response adjusted weights. Results in D-2 excluding 69 are significant and those in D-2a including 69 are not significant.

Tables D-3 and D-3a, also using fourth follow-up survey data, give results when the survey data are supplemented with SFA Pell recipient data and a longitudinal file made up of all sample members is used. The outcome variable is not standardized for EHSGY. As can be seen in these two models, results with and without 69 are significant.

Tables D-5 to D8 present data from responders to the third follow-up only with standardization for EHSGY for outcomes of postsecondary entrance evidence and presence on the aid file respectively. Tables D-7 and D-8 with third follow-up responders only look at presence on the aid file by +1 and show significance without project 69, but are not significant with project 69.

**Table D-1. Third follow-up *survey responders only, excludes Project 69, no SFA records, no standardization* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey only: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 77.8<br>pr-C = 72.2<br>Difference = 5.7**<br>psstu3<br>pweight: f3wgtsu | Variable name | Coef. | Linearized Std. Err. | T | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.369317 | 0.185317 | 1.99 | 0.053 | -0.00552 | 0.744156 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.38526 | 0.713454 | -0.54 | 0.592 | -1.82835 | 1.05784 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.51941 | 0.260255 | -2 | 0.053 | -1.04582 | 0.007009 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.55074 | 0.418589 | -1.32 | 0.196 | -1.39742 | 0.295935 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -2.93193 | 0.797048 | -3.68 | 0.001 | -4.54411 | -1.31975 |
| Clowoy (Low income only) | clowoy | 0.497022 | 0.33415 | 1.49 | 0.145 | -0.17886 | 1.172904 |
| Cfgenoy (First generation only) | cfgenoy | 0.502402 | 0.150072 | 3.35 | 0.002 | 0.198852 | 0.805951 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.594807 | 0.435353 | 1.37 | 0.18 | -0.28578 | 1.475391 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.033499 | 0.28969 | 0.12 | 0.909 | -0.55246 | 0.619453 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.27028 | 0.573049 | -0.47 | 0.64 | -1.42938 | 0.88882 |
| Cexdk (Baseline educational expectation was "don't know"---ref BA) | cexdk | -1.1039 | 0.133833 | -8.25 | 0 | -1.3746 | -0.8332 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -2.04626 | 0.362701 | -5.64 | 0 | -2.77989 | -1.31263 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.70324 | 0.222507 | -3.16 | 0.003 | -1.1533 | -0.25318 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.75898 | 0.153389 | -4.95 | 0 | -1.06924 | -0.44872 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.083534 | 0.153466 | 0.54 | 0.589 | -0.22688 | 0.393948 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.293215 | 0.197091 | 1.49 | 0.145 | -0.10544 | 0.691868 |
| Cothrac (Race was not Hispanic. Black, or White—ref Black) | cothrac | -0.30344 | 0.390095 | -0.78 | 0.441 | -1.09248 | 0.485601 |
| Chisp (Hispanic—ref Black) | chisp | -0.56072 | 0.334941 | -1.67 | 0.102 | -1.23821 | 0.11676 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.6727 | 0.1874 | -3.59 | 0.001 | -1.05176 | -0.29365 |
| Cfemale (Female) | cfemale | 0.482606 | 0.08674 | 5.56 | 0 | 0.307157 | 0.658055 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.418205 | 0.176585 | 2.37 | 0.023 | 0.061028 | 0.775383 |
| _cons | _cons | 1.136211 | 0.42492 | 2.67 | 0.011 | 0.276729 | 1.995692 |

NOTE: For this model only cases responding to the third follow-up survey were included. The same model (see D-2) is not significant when project 69 is included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 5 in body or report for additional note information. Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated April 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04.

**Table D-1a. Third follow-up *survey responders only, includes Project 69, no SFA records, no standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey only: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 76.4<br>pr-C = 75.4<br>Difference = 1.0 NS<br>psstu3 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.141031 | 0.268171 | 0.53 | 0.602 | -0.4014 | 0.683458 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.59776 | 0.663163 | -0.9 | 0.373 | -1.93913 | 0.743614 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.75743 | 0.266349 | -2.84 | 0.007 | -1.29617 | -0.21869 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.46113 | 0.426424 | -1.08 | 0.286 | -1.32365 | 0.401397 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -2.46243 | 0.934213 | -2.64 | 0.012 | -4.35205 | -0.57281 |
| Clowoy (Low income only) | clowoy | 0.479191 | 0.358532 | 1.34 | 0.189 | -0.24601 | 1.204391 |
| Cfgenoy (First generation only) | cfgenoy | 0.302686 | 0.208178 | 1.45 | 0.154 | -0.11839 | 0.723767 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.138236 | 0.559293 | 0.25 | 0.806 | -0.99304 | 1.269513 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.3085 | 0.376911 | -0.82 | 0.418 | -1.07088 | 0.45387 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.32094 | 0.582832 | -0.55 | 0.585 | -1.49983 | 0.857945 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -1.03927 | 0.112994 | -9.2 | 0 | -1.26782 | -0.81072 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -2.45028 | 0.527662 | -4.64 | 0 | -3.51758 | -1.38298 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -1.06833 | 0.287396 | -3.72 | 0.001 | -1.64964 | -0.48702 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.70579 | 0.135953 | -5.19 | 0 | -0.98078 | -0.4308 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.136026 | 0.126256 | 1.08 | 0.288 | -0.11935 | 0.391403 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.115257 | 0.196742 | 0.59 | 0.561 | -0.28269 | 0.513205 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.22205 | 0.375809 | -0.59 | 0.558 | -0.98219 | 0.538097 |
| Chisp (Hispanic—ref Black) | chisp | -0.17922 | 0.331225 | -0.54 | 0.592 | -0.84918 | 0.490748 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.56851 | 0.185982 | -3.06 | 0.004 | -0.94469 | -0.19232 |
| Cfemale (Female) | cfemale | 0.492856 | 0.088764 | 5.55 | 0 | 0.313315 | 0.672398 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.690804 | 0.251563 | 2.75 | 0.009 | 0.18197 | 1.199639 |
| _cons | _cons | 1.453551 | 0.460216 | 3.16 | 0.003 | 0.522675 | 2.384426 |

NOTE: For this model only cases responding to the third follow-up survey were included. The same model (see D-1) is significant when project 69 is excluded. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 5 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04.

**Table D-2. Fourth follow-up *survey responders only, excludes Project 69, no SFA records, no standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey only: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 83.3<br>pr-C = 79.0<br>Difference = 4.3***<br>ps_stud | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.337605 | 0.10175 | 3.32 | 0.002 | 0.131797 | 0.543413 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 1.30252 | 0.867437 | 1.5 | 0.141 | -0.45204 | 3.057077 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr99293 | 0.879648 | 0.535513 | 1.64 | 0.109 | -0.20353 | 1.962826 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 1.088942 | 0.812492 | 1.34 | 0.188 | -0.55448 | 2.732363 |
| Gr119293 (Grade 11 on baseline ref grade 9) | | | | | | | |
| Clowoy (Low income only) | clowoy | 0.633752 | 0.364168 | 1.74 | 0.09 | -0.10285 | 1.370352 |
| Cfgenoy (First generation only) | cfgenoy | 0.814161 | 0.213126 | 3.82 | 0 | 0.383073 | 1.24525 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.54535 | 0.767777 | -0.71 | 0.482 | -2.09833 | 1.007628 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.14976 | 0.495249 | -0.3 | 0.764 | -1.1515 | 0.851977 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.25548 | 0.864469 | -1.45 | 0.154 | -3.00403 | 0.493073 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.97584 | 0.228482 | -4.27 | 0 | -1.43799 | -0.5137 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.17703 | 0.385399 | -3.05 | 0.004 | -1.95657 | -0.39749 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.64872 | 0.258846 | -2.51 | 0.016 | -1.17229 | -0.12516 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.97565 | 0.27436 | -3.56 | 0.001 | -1.53059 | -0.4207 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.025812 | 0.317701 | 0.08 | 0.936 | -0.6168 | 0.668423 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.327316 | 0.210097 | 1.56 | 0.127 | -0.09764 | 0.752276 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.32522 | 0.396929 | 0.82 | 0.418 | -0.47765 | 1.128086 |
| Chisp (Hispanic—ref Black) | chisp | -0.44816 | 0.347514 | -1.29 | 0.205 | -1.15108 | 0.254751 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.72431 | 0.183085 | -3.96 | 0 | -1.09463 | -0.35398 |
| Cfemale (Female) | cfemale | 0.679909 | 0.129109 | 5.27 | 0 | 0.418763 | 0.941056 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.151738 | 0.235195 | 0.65 | 0.523 | -0.32399 | 0.627465 |
| _cons | _cons | 0.713349 | 0.251182 | 2.84 | 0.007 | 0.205287 | 1.221412 |

NOTE: Note this model is not significant when project 69 is included. For this model only cases responding to the fourth follow-up survey were included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 5 in text for complete note information; Number of strata = 27; Number of PSU = 66; Fourth Follow-up non-response adjusted weight used (f3wgtsu).
**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04.

**Table D-2a. Fourth follow-up *survey responders only, includes Project 69, no SFA records, no standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey only: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 80.4<br>pr-C = 80.1<br>Difference = .3NS | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.084736 | 0.215254 | 0.39 | 0.696 | -0.35066 | 0.520128 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 0.428275 | 1.020849 | 0.42 | 0.677 | -1.63659 | 2.493137 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.95684 | 0.402589 | -2.38 | 0.022 | -1.77115 | -0.14252 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 0.383448 | 0.636128 | 0.6 | 0.55 | -0.90324 | 1.670137 |
| Gr119293 (Grade 11 on baseline ref grade 9) | | | | | | | |
| Clowoy (Low income only) | clowoy | 0.593685 | 0.344787 | 1.72 | 0.093 | -0.10371 | 1.291083 |
| Cfgenoy (First generation only) | cfgenoy | 0.681439 | 0.149827 | 4.55 | 0 | 0.378385 | 0.984493 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -1.01867 | 0.809222 | -1.26 | 0.216 | -2.65548 | 0.618132 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.35362 | 0.382688 | -0.92 | 0.361 | -1.12768 | 0.420438 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.26615 | 0.817749 | -1.55 | 0.13 | -2.9202 | 0.387908 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.99535 | 0.18506 | -5.38 | 0 | -1.36967 | -0.62103 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.11906 | 0.298769 | -3.75 | 0.001 | -1.72338 | -0.51474 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.19011 | 0.49659 | -0.38 | 0.704 | -1.19455 | 0.814342 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.65298 | 0.375402 | -1.74 | 0.09 | -1.4123 | 0.106341 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.243658 | 0.311587 | 0.78 | 0.439 | -0.38659 | 0.873903 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.245596 | 0.167969 | 1.46 | 0.152 | -0.09415 | 0.585346 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.233535 | 0.377198 | 0.62 | 0.539 | -0.52942 | 0.996489 |
| Chisp (Hispanic—ref Black) | chisp | 0.115153 | 0.462875 | 0.25 | 0.805 | -0.8211 | 1.051406 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.48924 | 0.209367 | -2.34 | 0.025 | -0.91272 | -0.06576 |
| Cfemale (Female) | cfemale | 0.647159 | 0.137993 | 4.69 | 0 | 0.368041 | 0.926277 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.553313 | 0.337405 | 1.64 | 0.109 | -0.12915 | 1.235778 |
| _cons | _cons | 1.478867 | 0.45247 | 3.27 | 0.002 | 0.563661 | 2.394073 |

NOTE: This model is significant when project 69 is excluded. For this model only cases responding to the fourth follow-up survey were included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in text for complete note information; Number of strata = 27; Number of PSU = 66; Fourth follow-up survey non-response adjusted weight. (f3wgtsu).

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), US Department of Education: study conducted 1992-93 to 2003-04.

**Table D-3. Fourth follow-up *Longitudinal File, excludes project 69, uses SFA files, no standardization* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey or Pell Award Longitudinal File (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 76.9<br>pr-C = 68.6<br>Difference = 8.3****<br>newpost<br>v5bwgtp1 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.463618 | 0.099849 | 4.64 | 0 | 0.261655 | 0.665581 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 1.08042 | 0.58644 | 1.84 | 0.073 | -0.10577 | 2.266607 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr99293 | 0.409761 | 0.217525 | 1.88 | 0.067 | -0.03023 | 0.849747 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 0.28351 | 0.604111 | 0.47 | 0.641 | -0.93842 | 1.50544 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -1.91973 | 1.068992 | -1.8 | 0.08 | -4.08197 | 0.242508 |
| Clowoy (Low income only) | clowoy | 0.22122 | 0.243387 | 0.91 | 0.369 | -0.27108 | 0.713517 |
| Cfgenoy (First generation only) | cfgenoy | 0.502117 | 0.167035 | 3.01 | 0.005 | 0.164257 | 0.839977 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.238785 | 0.557928 | 0.43 | 0.671 | -0.88973 | 1.367301 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | 0.126644 | 0.27933 | 0.45 | 0.653 | -0.43835 | 0.691642 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.16844 | 0.556986 | -2.1 | 0.042 | -2.29505 | -0.04183 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.75869 | 0.188105 | -4.03 | 0 | -1.13917 | -0.37821 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -0.82877 | 0.223662 | -3.71 | 0.001 | -1.28117 | -0.37637 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.84796 | 0.251484 | -3.37 | 0.002 | -1.35664 | -0.33929 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.61834 | 0.179248 | -3.45 | 0.001 | -0.98091 | -0.25578 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | -0.00142 | 0.19485 | -0.01 | 0.994 | -0.39555 | 0.392697 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.247364 | 0.167452 | 1.48 | 0.148 | -0.09134 | 0.586067 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.03242 | 0.330711 | -0.1 | 0.922 | -0.70134 | 0.636507 |
| Chisp (Hispanic—ref Black) | chisp | -0.41865 | 0.338453 | -1.24 | 0.224 | -1.10323 | 0.265937 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.55126 | 0.136633 | -4.03 | 0 | -0.82763 | -0.2749 |
| Cfemale (Female) | cfemale | 0.708592 | 0.091058 | 7.78 | 0 | 0.524409 | 0.892774 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.073842 | 0.173554 | 0.43 | 0.673 | -0.2772 | 0.424889 |
| _cons | 0.302325 | 0.156342 | 1.93 | 0.06 | -0.01391 | 0.618557 | 0.302325 |

NOTE: This model is significant with and without project 69. Longitudinal file of all sample members; no standardization. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in text for additional information; Number of strata = 27; Number of PSU = 66; Uses poststratified weight --v5bwgtp1

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04.

**Table D-3a. Fourth follow-up *Longitudinal File, includes project 69, uses SFA files, no standardization*, Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey or Pell Award Longitudinal File (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 74.6<br>pr-C = 69.3<br>Difference = 5.3**<br>newpost<br>v5bwgtp1 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.32958 | 0.145134 | 2.27 | 0.029 | 0.036019 | 0.623141 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | 1.201728 | 0.601639 | 2 | 0.053 | -0.0152 | 2.418659 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr99293 | 0.614978 | 0.256851 | 2.39 | 0.022 | 0.095448 | 1.134508 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 1.581075 | 1.178788 | 1.34 | 0.188 | -0.80325 | 3.965398 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -0.55819 | 1.482307 | -0.38 | 0.709 | -3.55645 | 2.440055 |
| Clowoy (Low income only) | clowoy | 0.192644 | 0.225623 | 0.85 | 0.398 | -0.26372 | 0.64901 |
| Cfgenoy (First generation only) | cfgenoy | 0.586295 | 0.14419 | 4.07 | 0 | 0.294643 | 0.877946 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -1.13846 | 1.183651 | -0.96 | 0.342 | -3.53262 | 1.255705 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.23374 | 0.352961 | -0.66 | 0.512 | -0.94767 | 0.480189 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -1.20102 | 0.536453 | -2.24 | 0.031 | -2.28609 | -0.11594 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.76208 | 0.152513 | -5 | 0 | -1.07057 | -0.4536 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -0.78848 | 0.174431 | -4.52 | 0 | -1.1413 | -0.43566 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.9222 | 0.176962 | -5.21 | 0 | -1.28014 | -0.56426 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.41154 | 0.212039 | -1.94 | 0.06 | -0.84043 | 0.017345 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.130265 | 0.176729 | 0.74 | 0.465 | -0.2272 | 0.487732 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.009553 | 0.209893 | 0.05 | 0.964 | -0.415 | 0.434102 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | 0.033766 | 0.303158 | 0.11 | 0.912 | -0.57943 | 0.646961 |
| Chisp (Hispanic—ref Black) | chisp | -0.22797 | 0.240451 | -0.95 | 0.349 | -0.71433 | 0.258387 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.37358 | 0.144243 | -2.59 | 0.013 | -0.66533 | -0.08182 |
| Cfemale (Female) | cfemale | 0.794426 | 0.109682 | 7.24 | 0 | 0.572573 | 1.016279 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.538958 | 0.369007 | 1.46 | 0.152 | -0.20743 | 1.285344 |
| _cons | _cons | 0.067727 | 0.178514 | 0.38 | 0.706 | -0.29335 | 0.428806 |

NOTE: This model is significant with and without project 69. Longitudinal file of all sample members; no standardization. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in text for additional information; Number of strata = 28; Number of PSU = 67; Uses poststratified weight --v5bwgtp1

**SOURCE**: Data tabulated January 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education.

**Table D-4. Third follow-up survey, *survey responders only, excludes Project 69, includes standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey or SFA file by +1 of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 77.8<br>pr-C = 70.0<br>Difference = 7.8**** --Keyne2 standardized by EHSGY | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | ffutc | 0.471848 | 0.145736 | 3.24 | 0.002 | 0.177069 | 0.766627 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.20484 | 1.029489 | -0.2 | 0.843 | -2.28718 | 1.877502 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.40037 | 0.238112 | -1.68 | 0.101 | -0.882 | 0.081255 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.7047 | 0.478781 | -1.47 | 0.149 | -1.67312 | 0.263728 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -2.46541 | 0.978341 | -2.52 | 0.016 | -4.44429 | -0.48653 |
| Clowoy (Low income only) | clowoy | 0.362786 | 0.317251 | 1.14 | 0.26 | -0.27892 | 1.004488 |
| Cfgenoy (First generation only) | cfgenoy | 0.383814 | 0.245482 | 1.56 | 0.126 | -0.11272 | 0.880347 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.193708 | 0.575234 | 0.34 | 0.738 | -0.96981 | 1.357227 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.14841 | 0.279921 | -0.53 | 0.599 | -0.7146 | 0.417786 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.37578 | 0.777055 | -0.48 | 0.631 | -1.94752 | 1.195961 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.9754 | 0.165755 | -5.88 | 0 | -1.31067 | -0.64012 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -1.54376 | 0.451226 | -3.42 | 0.001 | -2.45645 | -0.63107 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.87067 | 0.265627 | -3.28 | 0.002 | -1.40795 | -0.33339 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.6131 | 0.152911 | -4.01 | 0 | -0.92239 | -0.30381 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.166976 | 0.141145 | 1.18 | 0.244 | -0.11852 | 0.452469 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.292371 | 0.153158 | 1.91 | 0.064 | -0.01742 | 0.602162 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.27156 | 0.327282 | -0.83 | 0.412 | -0.93355 | 0.390427 |
| Chisp (Hispanic—ref Black) | chisp | -0.46167 | 0.337168 | -1.37 | 0.179 | -1.14366 | 0.220311 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.65014 | 0.197479 | -3.29 | 0.002 | -1.04958 | -0.2507 |
| Cfemale (Female) | cfemale | 0.532785 | 0.105862 | 5.03 | 0 | 0.318658 | 0.746912 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.376998 | 0.210911 | 1.79 | 0.082 | -0.04961 | 0.803606 |
| _cons | _cons | 1.028935 | 0.411984 | 2.5 | 0.017 | 0.195618 | 1.862251 |

NOTE: For this model only cases responding to the Third Follow-up survey were included. Outcome variable is standardized. The same model (see D-5a) is marginally significant when project 69 is included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in body or report for additional note information. Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; Third Follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated May 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93-2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table D-4a. Third follow-up survey, *survey responders only, includes Project 69, includes standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of having evidence of entering postsecondary from survey or SFA file of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 75.9<br>pr-C = 71.4<br>Difference = 4.6*--Keyne2 standardized by EHSGY | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | Ffutc | 0.318136 | 0.17791 | 1.79 | 0.082 | -0.04172 | 0.677992 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.56832 | 0.947441 | -0.6 | 0.552 | -2.48471 | 1.348057 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.78901 | 0.382443 | -2.06 | 0.046 | -1.56257 | -0.01544 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.67931 | 0.48134 | -1.41 | 0.166 | -1.65292 | 0.294287 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -2.2214 | 1.02137 | -2.17 | 0.036 | -4.28731 | -0.15548 |
| Clowoy (Low income only) | Clowoy | 0.378319 | 0.343516 | 1.1 | 0.278 | -0.31651 | 1.073144 |
| Cfgenoy (First generation only) | Cfgenoy | 0.275751 | 0.222008 | 1.24 | 0.222 | -0.1733 | 0.724804 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.29124 | 0.686345 | -0.42 | 0.674 | -1.6795 | 1.097028 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.50922 | 0.399144 | -1.28 | 0.21 | -1.31656 | 0.298126 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.33268 | 0.734281 | -0.45 | 0.653 | -1.8179 | 1.152547 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.87859 | 0.166817 | -5.27 | 0 | -1.21601 | -0.54117 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -2.0257 | 0.624246 | -3.25 | 0.002 | -3.28836 | -0.76305 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -1.00957 | 0.206954 | -4.88 | 0 | -1.42817 | -0.59097 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.71782 | 0.157606 | -4.55 | 0 | -1.03661 | -0.39903 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.074933 | 0.155035 | 0.48 | 0.632 | -0.23865 | 0.388519 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | 0.027615 | 0.23187 | 0.12 | 0.906 | -0.44139 | 0.496615 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | Cothrac | -0.21328 | 0.329648 | -0.65 | 0.521 | -0.88006 | 0.453496 |
| Chisp (Hispanic—ref Black) | Chisp | -0.32536 | 0.219513 | -1.48 | 0.146 | -0.76937 | 0.118642 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.56228 | 0.191528 | -2.94 | 0.006 | -0.94968 | -0.17488 |
| Cfemale (Female) | Cfemale | 0.547214 | 0.092548 | 5.91 | 0 | 0.360018 | 0.734409 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.44684 | 0.160402 | 2.79 | 0.008 | 0.122397 | 0.771284 |
| _cons | _cons | 1.47975 | 0.527458 | 2.81 | 0.008 | 0.412865 | 2.546635 |

NOTE: For this model only cases responding to the third follow-up survey were included. Outcome variable is standardized. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated May 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table D-5. Third follow-up, *survey responders only, excludes Project 69, includes standardization,* Instrumental Variables regression (TOT) results for dependent variable of having evidence of entering postsecondary from survey or SFA file of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 78.2<br>pr-C = 65.6<br>Difference = 12.6*** Keyne2 standardized by EHSGY | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| xnewgp (evidence had some participation in UB or UBMS) | Xnewgp | 0.12093 | 0.041916 | 2.89 | 0.006 | 0.036147 | 0.205712 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.05136 | 0.163701 | -0.31 | 0.755 | -0.38248 | 0.279756 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.06701 | 0.044185 | -1.52 | 0.137 | -0.15638 | 0.02236 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.11024 | 0.097094 | -1.14 | 0.263 | -0.30663 | 0.086156 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -0.37186 | 0.164149 | -2.27 | 0.029 | -0.70388 | -0.03984 |
| Clowoy (Low income only) | Clowoy | 0.058254 | 0.050461 | 1.15 | 0.255 | -0.04381 | 0.16032 |
| Cfgenoy (First generation only) | Cfgenoy | 0.066584 | 0.044515 | 1.5 | 0.143 | -0.02346 | 0.156623 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.030226 | 0.116421 | 0.26 | 0.797 | -0.20526 | 0.26571 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.01582 | 0.051378 | -0.31 | 0.76 | -0.11974 | 0.088105 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.04571 | 0.121014 | -0.38 | 0.708 | -0.29049 | 0.199061 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.21047 | 0.040112 | -5.25 | 0 | -0.29161 | -0.12934 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -0.3324 | 0.104278 | -3.19 | 0.003 | -0.54332 | -0.12148 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.19009 | 0.061876 | -3.07 | 0.004 | -0.31525 | -0.06494 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.12471 | 0.031524 | -3.96 | 0 | -0.18847 | -0.06095 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.026265 | 0.02443 | 1.08 | 0.289 | -0.02315 | 0.07568 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | 0.043174 | 0.023858 | 1.81 | 0.078 | -0.00508 | 0.091431 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | Cothrac | -0.04565 | 0.061107 | -0.75 | 0.46 | -0.16925 | 0.077949 |
| Chisp (Hispanic—ref Black) | Chisp | -0.08405 | 0.066721 | -1.26 | 0.215 | -0.219 | 0.050908 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.11131 | 0.039783 | -2.8 | 0.008 | -0.19178 | -0.03085 |
| Cfemale (Female) | Cfemale | 0.105072 | 0.020994 | 5 | 0 | 0.062607 | 0.147537 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.057645 | 0.035567 | 1.62 | 0.113 | -0.0143 | 0.129587 |
| _cons | _cons | 0.695867 | 0.08441 | 8.24 | 0 | 0.525132 | 0.866603 |

NOTE: For this model only cases responding to the third follow-up survey were included. Outcome variable is standardized. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third Follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated May 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table D-5a. Third follow-up survey, *survey responders only, includes Project 69, includes standardization,* Instrumental Variables regression (TOT) results for dependent variable of having evidence of entering postsecondary from survey or SFA file of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = .76.0<br>pr-C = .68.2<br>Difference = 8.2 NS .11 Keyne2 standardized by EHSGY | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| xnewgp (evidence had some participation in UB or UBMS) | xnewgp | 0.079566 | 0.049383 | 1.61 | 0.115 | -0.02032 | 0.179453 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.1101 | 0.148731 | -0.74 | 0.464 | -0.41093 | 0.19074 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.13549 | 0.049163 | -2.76 | 0.009 | -0.23493 | -0.03605 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.11446 | 0.096604 | -1.18 | 0.243 | -0.30986 | 0.080939 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -0.37385 | 0.164861 | -2.27 | 0.029 | -0.70731 | -0.04039 |
| Clowoy (Low income only) | clowoy | 0.060933 | 0.05433 | 1.12 | 0.269 | -0.04896 | 0.170826 |
| Cfgenoy (First generation only) | cfgenoy | 0.046977 | 0.039249 | 1.2 | 0.239 | -0.03241 | 0.126365 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.05445 | 0.121787 | -0.45 | 0.657 | -0.30079 | 0.191882 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.08505 | 0.055154 | -1.54 | 0.131 | -0.19661 | 0.026505 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.04671 | 0.116594 | -0.4 | 0.691 | -0.28255 | 0.189119 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | cexdk | -0.18175 | 0.043381 | -4.19 | 0 | -0.2695 | -0.09401 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | cexhs | -0.44227 | 0.13162 | -3.36 | 0.002 | -0.70849 | -0.17604 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.20448 | 0.044165 | -4.63 | 0 | -0.29381 | -0.11515 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.14424 | 0.02802 | -5.15 | 0 | -0.20091 | -0.08756 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | cexma | 0.014779 | 0.026204 | 0.56 | 0.576 | -0.03822 | 0.067781 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | cexphd | 0.003308 | 0.035543 | 0.09 | 0.926 | -0.06858 | 0.0752 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | cothrac | -0.03762 | 0.063098 | -0.6 | 0.554 | -0.16525 | 0.090009 |
| Chisp (Hispanic—ref Black) | chisp | -0.0598 | 0.04181 | -1.43 | 0.161 | -0.14437 | 0.024766 |
| Cwhite (Race was White, not Hispanic—ref Black) | cwhite | -0.10072 | 0.039535 | -2.55 | 0.015 | -0.18069 | -0.02076 |
| Cfemale (Female) | cfemale | 0.108196 | 0.019756 | 5.48 | 0 | 0.068236 | 0.148156 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | parbefor | 0.068257 | 0.026527 | 2.57 | 0.014 | 0.0146 | 0.121913 |
| _cons | _cons | 0.787128 | 0.085408 | 9.22 | 0 | 0.614374 | 0.959882 |

NOTE: For this model only cases responding to the third follow-up survey were included. Outcome variable is standardized. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated May 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table D-6. Third follow-up, *survey responders only, includes Project 69,* Instrumental Variables regression (TOT) results for dependent variable of having evidence of entering postsecondary from survey no standardization: National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| xb-T = 75.4<br>xb-C =. 71.7.<br>Difference = 3.7 NS<br>Pweight = f3wgtsu<br>Psstu3 | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| xnewgp (evidence had some participation in UB or UBMS) | Xnewgp | 0.032015 | 0.067863 | 0.47 | 0.64 | -0.10525 | 0.169281 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.0986 | 0.097035 | -1.02 | 0.316 | -0.29488 | 0.097668 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.1198 | 0.022705 | -5.28 | 0 | -0.16573 | -0.07388 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.07644 | 0.072907 | -1.05 | 0.301 | -0.22391 | 0.071023 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -0.44269 | 0.122826 | -3.6 | 0.001 | -0.69113 | -0.19425 |
| Clowoy (Low income only) | Clowoy | 0.07375 | 0.050999 | 1.45 | 0.156 | -0.02941 | 0.176905 |
| Cfgenoy (First generation only) | Cfgenoy | 0.049848 | 0.035698 | 1.4 | 0.17 | -0.02236 | 0.122053 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | 0.036368 | 0.082454 | 0.44 | 0.662 | -0.13041 | 0.203148 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.0463 | 0.044216 | -1.05 | 0.302 | -0.13573 | 0.043138 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.05032 | 0.091234 | -0.55 | 0.584 | -0.23486 | 0.134213 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.21359 | 0.029249 | -7.3 | 0 | -0.27276 | -0.15443 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -0.51924 | 0.100293 | -5.18 | 0 | -0.7221 | -0.31637 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.21884 | 0.062181 | -3.52 | 0.001 | -0.34461 | -0.09307 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.13191 | 0.029941 | -4.41 | 0 | -0.19247 | -0.07135 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.021932 | 0.02081 | 1.05 | 0.298 | -0.02016 | 0.064023 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | 0.016635 | 0.029403 | 0.57 | 0.575 | -0.04284 | 0.076108 |
| Cothrac (Race was not Hispanic or Black or White—ref Black) | Cothrac | -0.03536 | 0.07011 | -0.5 | 0.617 | -0.17717 | 0.106448 |
| Chisp (Hispanic—ref Black) | Chisp | -0.03149 | 0.057064 | -0.55 | 0.584 | -0.14691 | 0.083933 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.10203 | 0.035308 | -2.89 | 0.006 | -0.17345 | -0.03061 |
| Cfemale (Female) | Cfemale | 0.090886 | 0.021805 | 4.17 | 0 | 0.046781 | 0.13499 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.104119 | 0.0359 | 2.9 | 0.006 | 0.031505 | 0.176734 |
| _cons | _cons | 0.786003 | 0.071313 | 11.02 | 0 | 0.641759 | 0.930247 |

NOTE: For this model only cases responding to the third follow-up survey were included. Outcome variable is not standardized. The same model is significant when project 69 is excluded. Ref = left out reference in dummy variable sequence. See table 5 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated September 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04.

**Table D-7. Third follow-up *survey responders only, excludes Project 69, includes standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of appearing on federal financial aid file as applicant within +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 67. 2<br>pr-C = 60.7<br>Difference =6.5** --kaidhs standardized by EHSGY | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | Ffutc | 0.335097 | 0.131333 | 2.55 | 0.015 | 0.069451 | 0.600744 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.62925 | 0.847138 | -0.74 | 0.462 | -2.34275 | 1.084246 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -0.62684 | 0.247231 | -2.54 | 0.015 | -1.12691 | -0.12676 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | -0.2284 | 0.584013 | -0.39 | 0.698 | -1.40968 | 0.952875 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -1.97727 | 1.054882 | -1.87 | 0.068 | -4.11097 | 0.15643 |
| Clowoy. (Low income only) | Clowoy | 0.076876 | 0.252053 | 0.3 | 0.762 | -0.43295 | 0.5867 |
| Cfgenoy (First generation only) | Cfgenoy | 0.182671 | 0.23071 | 0.79 | 0.433 | -0.28398 | 0.649325 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.04901 | 0.667762 | -0.07 | 0.942 | -1.39969 | 1.301666 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.29504 | 0.275853 | -1.07 | 0.291 | -0.853 | 0.262927 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | -0.11524 | 0.590867 | -0.2 | 0.846 | -1.31038 | 1.079904 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.94322 | 0.223889 | -4.21 | 0 | -1.39608 | -0.49037 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -1.46153 | 0.309387 | -4.72 | 0 | -2.08732 | -0.83573 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -0.61096 | 0.2129 | -2.87 | 0.007 | -1.04159 | -0.18033 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -0.74734 | 0.123378 | -6.06 | 0 | -0.9969 | -0.49779 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.152594 | 0.122309 | 1.25 | 0.22 | -0.0948 | 0.399988 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | 0.133778 | 0.092662 | 1.44 | 0.157 | -0.05365 | 0.321205 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | Cothrac | -0.06428 | 0.26635 | -0.24 | 0.811 | -0.60302 | 0.474468 |
| Chisp (Hispanic—ref Black) | Chisp | -0.41505 | 0.263606 | -1.57 | 0.123 | -0.94824 | 0.118145 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.60977 | 0.199767 | -3.05 | 0.004 | -1.01384 | -0.20571 |
| Cfemale (Female) | Cfemale | 0.523694 | 0.129561 | 4.04 | 0 | 0.261633 | 0.785755 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.200495 | 0.114352 | 1.75 | 0.087 | -0.0308 | 0.431794 |
| _cons | _cons | 0.869016 | 0.372415 | 2.33 | 0.025 | 0.115735 | 1.622297 |

NOTE: For this model only cases responding to the third follow-up survey were included. Outcome variable is standardized. The same model (see D-6a) is not significant when project 69 is included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 5 in body or report for additional note information. Number of strata (v5no69st) = 27; Number of PSU (wprojid) = 66; Third follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated May 2008 using: National Evaluation of Upward Bound data files, study sponsored by the Policy and Program Studies Services (PPSS), of the Office of Planning, Evaluation and Policy Development (OPEPD), U.S. Department of Education: study conducted 1992-93 to 2003-04; federal Student Financial Aid (SFA) files 1994-95 to 2003-04.

**Table D-8. Third follow-up *survey responders only, not longitudinal file, includes Project 69, includes standardization,* Intent to Treat (ITT) logistic regression results for dependent variable of appearing on federal financial aid file as applicant within +1 (18 months) of expected high school graduation year (EHSGY): National Evaluation of Upward Bound, study conducted 1992-93 to 2003-04**

| pr-T = 62.0<br>pr-C = 59.1<br>Difference = 2.8 NS<br>kaidhs | Variable name | Coef. | Linearized Std. Err. | t | P>\|t\| | 95% Confidence Interval | |
|---|---|---|---|---|---|---|---|
| FFUTC (random assigned to treatment) | Ffutc | 0.201285 | 0.15696 | 1.28 | 0.207 | -0.1162 | 0.518766 |
| Gr79293 (Grade 7 on baseline ref grade 9) | gr79293 | -0.99231 | 0.683285 | -1.45 | 0.154 | -2.37438 | 0.389767 |
| Gr89293 (Grade 8 on baseline ref grade 9) | gr89293 | -1.1552 | 0.502105 | -2.3 | 0.027 | -2.1708 | -0.1396 |
| Gr109293 (Grade 10 on baseline ref grade 9) | gr109293 | 0.041714 | 0.619753 | 0.07 | 0.947 | -1.21185 | 1.295282 |
| Gr119293 (Grade 11 on baseline ref grade 9) | gr119293 | -1.45371 | 1.128657 | -1.29 | 0.205 | -3.73664 | 0.829212 |
| Clowoy (Low income only) | Clowoy | 0.268405 | 0.305969 | 0.88 | 0.386 | -0.35048 | 0.887286 |
| Cfgenoy (First generation only) | Cfgenoy | 0.176501 | 0.176561 | 1 | 0.324 | -0.18063 | 0.533629 |
| C11gssf (Grade was 11 on student selection form—ref grade 9) | c11gssf | -0.79206 | 0.880256 | -0.9 | 0.374 | -2.57255 | 0.988424 |
| C10gssf (Grade was 10 on student selection form—ref grade 9) | c10gssf | -0.8853 | 0.549069 | -1.61 | 0.115 | -1.99589 | 0.225298 |
| C8gssfm (Grade was 8 on student selection form—ref grade 9) | c8gssfm | 0.067778 | 0.495135 | 0.14 | 0.892 | -0.93373 | 1.069283 |
| Cexdk (Baseline educational expectation was "don't know" ---ref BA) | Cexdk | -0.80728 | 0.22559 | -3.58 | 0.001 | -1.26358 | -0.35098 |
| Cexhs (Baseline educational expectation was high school only---ref BA) | Cexhs | -1.76048 | 0.435647 | -4.04 | 0 | -2.64166 | -0.87931 |
| Cex13v (Baseline educational expectation was vocational---ref BA) | cex13v | -1.0296 | 0.312666 | -3.29 | 0.002 | -1.66202 | -0.39717 |
| cex14aa (Baseline educational expectation was two-year---ref BA) | cex14aa | -1.00794 | 0.260691 | -3.87 | 0 | -1.53524 | -0.48065 |
| Cexma (Baseline educational expectation was Masters Degree---refer BA) | Cexma | 0.18785 | 0.094023 | 2 | 0.053 | -0.00233 | 0.37803 |
| Cexphd (Baseline educational expectation was PhD---ref BA) | Cexphd | -0.16609 | 0.226026 | -0.73 | 0.467 | -0.62327 | 0.291092 |
| Cothrac (Race was not Hispanic, Black, or White—ref Black) | Cothrac | 0.37531 | 0.391083 | 0.96 | 0.343 | -0.41573 | 1.166349 |
| Chisp (Hispanic—ref Black) | Chisp | -0.08086 | 0.254593 | -0.32 | 0.752 | -0.59582 | 0.434104 |
| Cwhite (Race was White, not Hispanic—ref Black) | Cwhite | -0.18963 | 0.315171 | -0.6 | 0.551 | -0.82712 | 0.447861 |
| Cfemale (Female) | Cfemale | 0.650257 | 0.197571 | 3.29 | 0.002 | 0.250632 | 1.049883 |
| Parbefor (Reported participated in other pre-college supplemental services before random assignment) | Parbefor | 0.344947 | 0.130587 | 2.64 | 0.012 | 0.080809 | 0.609084 |
| _cons | _cons | 1.012604 | 0.38979 | 2.6 | 0.013 | 0.22418 | 1.801029 |

NOTE: For this model only cases responding to the third Follow-up survey were included. Outcome variable is standardized. The same model (see D-6) is significant when project 69 is included. SFA = Student Financial Aid file records; Ref = left out reference in dummy variable sequence. See table 6 in body or report for additional note information. Number of strata (wprstco) = 28; Number of PSU (wprojid) = 67; Third Follow-up non-response adjusted weight (f3wgtsu).

**SOURCE**: Data tabulated (May 2008) by Policy and Program Studies Services (PPSS) using data from the, National Evaluation of Upward Bound, as applicable, study files baseline through fifth follow-up; Federal Student Financial Aid (SFA) files: 1994-95 to 2003-04.

WhatWorks

---

From:WhatWorks
Sent:9 Apr 2014 14:04:41 +0000
To:'margaret.cahalan@pellinstitute.org'
Subject:RE: WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review (WWC 4643)
Hello,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

---

**From:** Margaret Cahalan [mailto:margaret.cahalan@pellinstitute.org]
**Sent:** Monday, April 07, 2014 10:44 AM
**To:** Info@whatworks.ed.gov
**Cc:** (b)(6)
**Subject:** WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review

Dear What Works Clearinghouse,

We have attached 5 files providing our formal request for a Quality Review of the WWC ratings listed below.

- Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular **Upward Bound**: Results from the third follow-up data collection*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbo ok
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

- Seftor, N. S., Mamun, A., & Schirm, A. (2009). *The impacts of regular **Upward Bound** on Qpostsecondary outcomes 7–9 years after scheduled high school graduation*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbook
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

The attached documents describe and document 10 specific concerns and violations of WWC Standards and of NCES and general statistical and evaluation research standards. We look forward to a timely and fair review of the concerns we have raised. If you have any questions regarding this submission, Dr. Goodwin and I would be happy to address these questions. Our contact information is below.

**Margaret Cahalan**
Margaret.cahalan@pellinstitute.org
202-347-7430 ex 207
(b)(6) (c)

**David Goodwin**
(b)(6)

Thank you for your consideration, Kind regards,
Maggie Cahalan

Margaret Cahalan, Ph.D, Vice President for Research, Director Pell Institute,. email Margaret.cahalan@pellinstitute.org
Co-Principal Investigator i-3 grant Using Data to Inform College Access Programming
Council for Opportunity in Education (COE) www.coenet.us;
THE PELL INSTITUTE for the Study of Opportunity in Higher Education www.pellinstitute.org
1025 Vermont Ave., NW, Suite 1020, Washington, DC 20005, P: 202-347-7430 ex 207--C: (b)(6)

WhatWorks

---

From:WhatWorks
Sent:10 Apr 2014 20:26:42 +0000
To:'margaret.cahalan@pellinstitute.org'
Subject:RE: WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review (WWC 4643)
Hello,

The WWC quality review team is reviewing your email and will prepare a written response. The quality review team responds to concerns raised by study authors, curriculum developers, or other relevant parties about WWC reviews published on our website. These quality reviews are undertaken when concerned parties present evidence that a WWC review might be inaccurate. When a quality review is conducted, a researcher who was not involved in the initial review undertakes an independent assessment of the study in question. The researcher also investigates the procedures used and decisions made during the original review of the study. If a quality review concludes that the original review contained errors, a revision will be published. These quality reviews are one of tools used to ensure that the standards established by the Institute of Education Sciences are upheld on every review conducted by the WWC.

Thank you,

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

---

**From:** WhatWorks
**Sent:** Wednesday, April 09, 2014 9:05 AM
**To:** 'margaret.cahalan@pellinstitute.org'
**Subject:** RE: WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review (WWC 4643)

Hello,

Thank you for contacting the What Works Clearinghouse (WWC). We have received your email below. WWC staff are reviewing your request and will prepare a response.

What Works Clearinghouse

The What Works Clearinghouse was established by the U.S. Department of Education's Institute of Education Sciences to provide educators, policymakers, researchers, and the public with a central and trusted source of scientific evidence of what works in education. For more information, please visit http://ies.ed.gov/ncee/wwc/.

---

**From:** Margaret Cahalan [mailto:margaret.cahalan@pellinstitute.org]
**Sent:** Monday, April 07, 2014 10:44 AM
**To:** Info@whatworks.ed.gov
**Cc:** (b)(6)

**Subject:** WWC Contact Reference number is 2086160509--National Evaluation of Upward Bound Reports--Request for Quality Review

Dear What Works Clearinghouse,

We have attached 5 files providing our formal request for a Quality Review of the WWC ratings listed below.

- Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular **Upward Bound**: Results from the third follow-up data collection*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbo ok
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

- Seftor, N. S., Mamun, A., & Schirm, A. (2009). *The impacts of regular **Upward Bound** on Qpostsecondary outcomes 7–9 years after scheduled high school graduation*. Princeton, NJ: Mathematica Policy Research.
  Rating: Meets evidence standards without reservations
  Reviewed using: WWC Procedures and Standards Handbook
  Reviewed in Practice Guide: Helping Students Navigate the Path to College: What High Schools Can Do

The attached documents describe and document 10 specific concerns and violations of WWC Standards and of NCES and general statistical and evaluation research standards. We look forward to a timely and fair review of the concerns we have raised. If you have any questions regarding this submission, Dr. Goodwin and I would be happy to address these questions. Our contact information is below.

**Margaret Cahalan**
Margaret.cahalan@pellinstitute.org
202-347-7430 ex 207
(b)(6) (c)

**David Goodwin**
(b)(6)

Thank you for your consideration, Kind regards,
Maggie Cahalan

Margaret Cahalan, Ph.D, Vice President for Research, Director Pell Institute, email Margaret.cahalan@pellinstitute.org
Co-Principal Investigator i-3 grant Using Data to Inform College Access Programming
Council for Opportunity in Education (COE) www.coenet.us;
THE PELL INSTITUTE for the Study of Opportunity in Higher Education www.pellinstitute.org

1025 Vermont Ave., NW, Suite 1020, Washington, DC 20005, P: 202-347-7430 ex 207--C: (b)(6)

May 16, 2014

Dear Drs. Cahalan and Goodwin,

Thank you for your inquiry concerning the What Works Clearinghouse (WWC) rating of the following two study reports for the "Helping Students Navigate the Path to College: What High Schools Can Do" Practice Guide:

Myers, D., Olsen, R., Seftor, N., Young, J., & Tuttle, C. (2004). *The impacts of regular* ***Upward Bound****: Results from the third follow-up data collection*. Princeton, NJ: Mathematica Policy Research.

Seftor, N. S., Mamun, A., & Schirm, A. (2009). *The impacts of regular* ***Upward Bound*** *on postsecondary outcomes 7–9 years after scheduled high school graduation*. Princeton, NJ: Mathematica Policy Research.

In response to your inquiry, we conducted an independent quality review to address the issues raised. The WWC Quality Review Team (QRT) responds to concerns raised about WWC reviews published on our website. When a quality review is conducted, researchers who were not involved in the initial review undertake an independent assessment of the study in question. The researchers also investigate the procedures used and decisions made during the original review of the study. These quality reviews are one of the tools used to ensure that the standards established by the Institute of Education Sciences (IES) are upheld for every review conducted by the WWC.

In your letter dated April 7, 2014, you raised a number of items related to the two study reports referenced above. The issues are quoted below as you described them:

1. "ED Information Guidelines and NCES Standards Require an Adequate Sample Design —Serious Unequal Weighing."
2. "NCES Standards and ED Guidelines Require a Representative Sample for Estimation of Averages."
3. "WWC Standards Require A Balance Between the Treatment and Control Group At Baseline on Factors Likely to Impact Outcomes. (Non-Equivalence of Treatment and Control Group)."
4. "WWC Standards Require that the Observed Impact be Attributable to the Intervention."
5. "WWC Standards Require Use of a Common Outcome Measure for Impact Estimation. Lack of Precision and Standardized Common Outcome Measures."

6. "AERA and ED Guidelines Require "Warranted" Conclusions."
7. "NCES Coverage Standards Require that the Data Sources Used Have Adequate and Non-Biased Coverage."
8. "AERA and ED Guidelines Require "Warranted" Conclusions-- Failure to acknowledge large positive results for BA for evenly matched treatment and control group."
9. "WWC Attribution Standards Specify that the Impact must be Attributable to the Intervention. --Control Group Contamination Issues are Not Acknowledged."
10. "The Mathematica UB reports lack transparency and violate the "stakeholders right to know" and the "sufficiency of the warrants" basic standards for evaluation research."

Five issues – 1, 2, 6, 7, and 8 – involve ED Information Guidelines, NCES standards, and AERA standards. These five issues refer to guidelines that are beyond the purview of the WWC. Therefore, we do not address these issues in the QRT review.

Issue 10 focuses on transparency. This issue is not covered in the WWC standards. Therefore, we do not comment on this issue as it relates to the WWC ratings of the *Upward Bound* study.

The quality review team did, however, investigate Issues 3, 4, 5, and 9, which were related to the WWC standards. The issues are listed below, numbered according to your request, along with the quality review team's response to each issue.

**ISSUE 3. "WWC Standards Require A Balance Between the Treatment and Control Group At Baseline on Factors Likely to Impact Outcomes. (Non-Equivalence of Treatment and Control Group)."**

**Response:** For randomized controlled trials with low attrition, WWC standards do not require demonstration of baseline equivalence (WWC Handbook v. 3.0)[1]. The quality review determined that in the *Upward Bound* study, overall and differential attrition are low for the following outcomes: high school credit accumulation, high school GPA, high school completion, post-secondary enrollment, post-secondary degree attainment, and financial aid application and receipt. Thus, the study is not required to demonstrate baseline equivalence for these outcomes, which meet WWC group design standards *without* reservations.

[1] According to WWC procedures, the Quality Review Team reviews studies according to the current version of the standards, which is 3.0. This study was originally reviewed for the "Helping Students Navigate the Path to College: What High Schools Can Do" Practice Guide which was published in 2009, and used WWC Standards version 1.0. The standards regarding attrition and baseline equivalence for RCTs have not changed.

For some study outcomes, specifically those related to postsecondary credit accumulation, the WWC determined that attrition is high (reported on p. 36 of Myers et al., 2004). For these outcomes, although the impact estimates control for some baseline covariates, a prior achievement measure is not included among the covariates, as is required to meet WWC standards (see p. 7 of Myers et al., 2004). Therefore, the outcomes related to postsecondary credit accumulation do not meet WWC group design standards.

In sum, according to the WWC Standards and Procedures Handbook, when outcomes within a study are assigned different ratings – as is the case with the *Upward Bound* study – the highest rating is recorded as the overall study rating. Therefore, because the *Upward Bound* study has outcomes that meet standards without reservations, the entire study receives that rating. On this basis, the quality review team concluded that the rating of meets standards without reservations is appropriate for the study.

**ISSUE 4. "WWC Standards Require that the Observed Impact be Attributable to the Intervention."**

**Response:** On this point, the Quality Review Team was unable to determine which of two WWC standards you referenced; explanations for both possibilities are provided below. First, the attribution standard for regression discontinuity designs requires that there be continuity in the outcome-forcing variable relationship. This standard does not apply to the *Upward Bound* study since it is a RCT and not regression discontinuity study. Second, WWC standards for group designs require that there be no evidence of confounding factors. That is, the intervention or comparison conditions cannot be perfectly aligned with a component of the study design or the circumstances under which the intervention was implemented. There was no evidence of a confounding factor in this study. The Quality Review Team concluded that the original assessment by the WWC was appropriate.

**ISSUE 5. "WWC Standards Require Use of a Common Outcome Measure for Impact Estimation. Lack of Precision and Standardized Common Outcome Measures."**

Response: To be eligible for review, an outcome must (a) demonstrate face validity and reliability, (b) not be overaligned with the intervention, and (c) be collected in the same manner for both intervention and comparison groups. The Quality Review Team concluded that the outcomes used to establish the rating for the *Upward Bound* study met all three of these requirements.

**ISSUE 9. "WWC Attribution Standards Specify that the Impact must be Attributable to the Intervention. --Control Group Contamination Issues are Not Acknowledged."**

**Response:** As above, the Quality Review Team was unable to determine which of two WWC standards you referenced. Therefore, two explanations are provided. The attribution standard that applies to regression discontinuity designs was not applied to the Upward Bound study, which used a randomized control trial design. Second, when reviewing studies, the WWC standards focus on design and analysis to determine whether the analysis provides a rigorous test of the intervention being examined. In this case, though the study authors acknowledge that some students in the comparison condition received similar services, the Quality Review Team found that conducting the analysis under the intent-to-treat framework – the authors' choice in the *Upward Bound* study – provided the most rigorous test of the intervention. Therefore, the Quality Review Team concluded that the original assessment by the WWC was appropriate.

In summary, after a thorough review of all issues relevant to the WWC, the Quality Review Team has determined that the WWC should retain the current rating of the *Upward Bound* study.

I hope that this letter has addressed your concerns. If you have other concerns, please do not hesitate to contact the WWC through info@whatworks.ed.gov. So that we can maintain a written record of key decisions, we ask that you submit questions to us in writing.

Kind regards,

(b)(6)

Sandra Jo Wilson, PhD
Principal Investigator, What Works Clearinghouse at Development Services Group, Inc.
Associate Director, Peabody Research Institute, Vanderbilt University

whatworks@dsgonline.com

---

From:whatworks@dsgonline.com
Sent:16 May 2014 13:47:59 +0000
To:margaret.cahalan@pellinstitute.org;(b)(6)
Cc:What Works;info@whatworks.ed.gov
Subject:Response to WWC Query
Attachments:Response to Cahalan and Goodwin_FINAL.pdf

Dear Drs. Cahalan and Goodwin,

Please find attached a response to your query to the WWC regarding the Upward Bound study. If you have additional questions, please let us know by contacting the help desk at info@whatworks.ed.gov.

Kind regards,
Sandra Wilson