## Appendix K Imputations

Following data imputations described here, variables were reviewed and revised, if necessary, to adjust for inconsistencies with other known data. Therefore, the "after imputation" distributions may differ from the final distributions in the restricted-use data file.

Appendix K: Imputations

As described in Section 5.4, data for 23 variables were imputed statistically, mostly using the weighted hot deck procedure. This appendix shows the imputation classes and sorting variables for all of the variables imputed by the hot deck approach, as well as the other imputation procedures that were used. As presented in table 5.4, the variables are listed in the order in which the imputations were performed.

```
(1) Variable Name:
Description:
Data Used in Imputations: Study respondents (61,767)
Number Missing: }343\mathrm{ (0.56%)
Imputation Classes: A cross-classification of
Student type }\mp@subsup{}{}{1
Undergraduate/graduate level
Dependency status
Student marital status
Fall attendance status
(2) Variable Name:
Description:
Data Used in Imputations:
Number Missing:
Imputation Classes:
Sorting Variables:
Hispanic indicator
Student type
(3) Variable Name:
Description:
CITIZEN2
Student citizenship
Data Used in Imputations: \(\quad\) Study respondents \((61,767)\)
Number Missing: \(\quad 2,408\) (3.90\%)
Imputation Classes: A cross-classification of
Federal student aid receipt status
Student type
Sorting Variables:
Institution control
Institution level of instruction
Race (simplified)
```

[^0]| Variable Name: | HISPANIC |
| :--- | :--- |
| Description: | Indicator of Hispanic ethnicity |
| Data Used in Imputations: | Study respondents $(61,767)$ |
| Number Missing: | $3,087(5.00 \%)$ |
| Imputation Classes: | A cross-classification of |
| $\quad$ OBE region |  |
| $\quad$ Federal student aid receipt status |  |
| $\quad$ Percent of Hispanics at institution (categorical) |  |
| Sorting Variable: |  |
| $\quad$ First name |  |

## Variable Name: <br> Description:

RACE
An intermediary variable allowing for a full racial pattern of all possible multiple-listings of races (31 possible values). This variable was formed from the variables that were individual race indicators: R2WHITE, R2BLACK, R2ASIAN, R2ISLAND, and R2INDIAN. After RACE was imputed, the variables R2WHITE, R2BLACK, R2ASIAN, R2ISLAND, and R2INDIAN were logically assigned from the values of RACE. Study respondents $(61,767)$

## $\begin{array}{ll}\text { Data Used in Imputations: } & \text { Study responde } \\ \text { Number Missing: } & 4,968(8.04 \%) .\end{array}$

Imputation Procedure:
These missing values were those for study respondents who had given no positive response for any of the five racial indicators. The assumption here is that respondents who marked at least one racial category with a positive indication have given a sufficiently complete self-profile even if other racial categories were left missing. Details for each of the 5 racial indicator variables are described below.
Imputation classes: A cross-classification of
Student type
Percent of blacks at institution (categorical)
Percent of Hispanics at institution (categorical)
Percent of Asian/Pacific islanders at institution (categorical)
Percent of American Indians at institution (categorical)
Sorting Variables:
Hispanic indicator
Percent of whites at institution (continuous)
OBE region
(5b) Variable Name: R2WHITE
Description:
White race indicator
Data Used in Imputations: $\quad$ Study respondents $(61,767)$
Number Missing: 5,005 (8.10\%)
Imputation Procedure: Logically imputed from the value of RACE: 1 if RACE had value of white; 0 otherwise
(5c) Variable Name:
Description:
R2BLACK
Black race indicator
Data Used in Imputations: Study respondents $(61,767)$
Number Missing: $\quad 5,147$ (8.33\%)
Imputation Procedure: Logically imputed from the value of RACE: 1 if RACE had value of black; 0 otherwise
(5d) Variable Name:
R2ASIAN

[^1]Description: Asian race indicator
Data Used in Imputations: $\quad$ Study respondents $(61,767)$
Number Missing:
Imputation Procedure:

5,178 (8.38\%)
Logically imputed from the value of RACE: 1 if RACE had value of Asian; 0 otherwise

Variable Name:
Description:
Data Used in Imputations:
Number Missing:
Imputation Procedure:

R2ISLAND
Pacific Islander race indicator
Study respondents $(61,767)$
5,178 (8.38\%)
Logically imputed from the value of RACE: 1 if RACE had value of Pacific Islander; 0 otherwise
Variable Name: R2INDIAN

Description:
American Indian race indicator
Data Used in Imputations: $\quad$ Study respondents $(61,767)$
Number Missing:
5,172 (8.37\%)
Imputation Procedure: Logically imputed from the value of RACE: 1 if RACE had value of American Indian; 0 otherwise
Variable Name:
Description:
Data Used in Imputations:
Number Missing:
Imputation Classes:
$\quad$ CPS record indicator
Fall enrollment status
$\quad$ Student type
$\quad$ Age (categorical) ${ }^{4}$
Sorting Variable:
$\quad$ Age

SMARITAL
Student marital status
Study respondents $(61,767)$
5,032 (8.15\%)
10 CHAID segments defined by cross-classifications of
CPS record indicator
Fall enrollment status
Student type
Age (categorical) ${ }^{4}$

## Sorting Variable:

Age
Variable Name:
Description:
ANYDEP
Dependents indicator
Data Used in Imputations: $\quad$ Study respondents $(61,767)$
Number Missing:
9,179 (14.86\%)
Imputation Classes:
8 CHAID segments defined by cross-classifications of
Black race indicator
Gender
Fall enrollment status
Student marital status
Hispanic indicator

[^2](8a)

| Variable Name: | DEPEND |
| :--- | :--- |
| Description: | Dependency status indicator (2 levels) |
| Data Used in Imputations: | Study respondents $(61,767)$ |
| Number Missing: | $3,969(6.43 \%)$ |

Imputation Procedure:
Missing values were imputed based on age, student marital status, and whether or not the respondent has any dependents. A person was considered an "Independent" if he/she had any dependents (ANYDEP=1), or if he/she was at least 24 years of age (AGE ge 24), or if he/she was married or separated (SMARITAL is (2 or 3 )), or if he/she was a graduate or first-professional student (STUTYPE5 is (2 or 3)). Otherwise, the student was considered to be a "Dependent."
(8b) Variable Name:
Description:
Data Used in Imputations: $\quad$ Study respondents $(61,767)$
Number Missing:
Imputation Procedure:
Missing values were imputed based on the values of the first dependency status indicator (DEPEND) and the indicator of any dependents (ANYDEP). If a study respondent has already been identified as independent by DEPEND, and he/she has dependents, then DEPEND2 will indicate whether or not he had any dependents.

## Variable Name:

Description:
Data Used in Imputations: $\quad$ Study respondents for whom FALL $>0(51,232)$
Number Missing:
Imputation Classes:

Federal aid receipt indicator
Student marital status
Sorting Variable:
Age
(10) Variable Name:

Description:
Data Used in Imputations:
Number Missing:
Imputation Classes:
Citizenship
Student type
Institution level of instruction
Sorting Variables:
Institution highest level of offering
Age
Race (simplified)

| Variable Name: | LOCALRES |
| :--- | :--- |
| Description: | Local residence |
| Data Used in Imputations: | Study respondents $(61,767)$ |
| Number Missing: | 10,704 (17.33\%) |
| Imputation Classes: | 8 CHAID segments defined from cross-classifications of |
| $\quad$ Dependency status indicator (2 levels) |  |
| Hispanic indicator |  |
| Fall attendance status |  |
| $\quad$ CPS record indicator |  |
| $\quad$ Dependents indicator |  |
| $\quad$ Citizenship |  |
| Student marital status |  |
| Sorting Variable: |  |
| Age |  |

(12) Variable Name:

Description:
Data Used in Imputations:
Number Missing:

NDEPEND
Number of dependents
Study respondents $(61,767)$
11,328 (18.34\%) Of study respondents with dependents (ANYDEP=1), there were 4,673 missing $(29.98 \%$ of 15,586$)$.
Of study respondents who reported no dependents
(ANYDEP=0), there were 6,655 missing NDEPEND values (14.41\% of 46,181).

## Imputation Procedure:

Hot deck imputation was implemented for all missing NDEPEND values for which ANYDEP=1. In this group, there were 853 respondents with an NDEPEND value of zero. Since this NDEPEND value was inconsistent with the ANYDEP value, it made these respondents inappropriate donors. Thus, they were eliminated from the donor base. Of the 46,181 study respondents for which ANYDEP $=0$, there were 6,655 with missing NDEPEND values. All of these were logically set to zero.
Imputation Classes: A cross-classification of
Student marital status
Age (categorical)
Gender
Student type (Graduate students were collapsed into a single group with first-professional students)
Sorting Variables:
Age
CPS record indicator
Institution level of instruction

| Variable Name: | PMARITAL |
| :--- | :--- |
| Description: | Parents' marital status |
| Data Used in Imputations: | Study respondents, dependents $(26,167)$ |
| Number Missing: | $3,582(13.69 \%)$ |
| Imputation Classes: | A cross-classification of |

Institution highest level of offering
Race ${ }^{5}$ (simplified)
Age (categorical)

## Sorting Variables:

Age
Race
(14) Variable Name:

PFAMNUM
Description:
Parent family size
Data Used in Imputations: Study respondents, dependents $(26,167)$
Number Missing: 3,582 (13.69\%)
Imputation Classes:
A cross-classification of
Parents' marital status
Institution highest level of offering
Race ${ }^{5}$ (simplified)

## Sorting Variables:

Age
Race

Variable Name:
Description:
Data Used in Imputations: Study respondents, dependents $(26,167)$
Number Missing:

## DEPINC

Parents' income
overall: 10,503 (40.14\%)
$1^{\text {st }}$ stage: $6,901(48.29 \%$ of 14,292 dependents reporting parents' income category)
$2^{\text {nd }}$ stage: 3,602 ( $18.96 \%$ of dependents not imputed in $1^{\text {st }}$ stage)
Imputation Classes, $1^{\text {st }}$ Stage: A cross-classification of
Parent income category from student reports (if reported)
Parent marital status
Imputation Classes, 2nd Stage: 9 CHAID segments defined by cross-classifications of Pell grant status
Parent marital status
Citizenship
Hispanic indicator
Parent family size

[^3]Sorting Variables:<br>Parents' highest education<br>Race

NOTE: The imputation for parent income was performed in two stages. The first stage imputed for students who reported their parents' income category but the actual parents' income amount was missing. The first stage used a cross-classification of parent income category and parent marital status as the imputation classes among students who reported their parents' income category. The second stage imputed for students who did not report both their parents' income category and actual parent's income. The second stage imputed the remaining missing values where several variables were used to define the imputation classes, including parent marital status, which was also used as an imputation class in the first stage. In both stages, parents' highest education and race were used as the sorting variables.

| Variable Name: | HSGRADYY |
| :--- | :--- |
| Description: | High school graduation year |
| Data Used in Imputations: | Study respondents, high school degree or certificate $(61,058)$ |
| Number Missing: | $8,416(13.78 \%)$ |
| Imputation Procedure: |  |

It was assumed HSGRADYY was not missing for respondents who had indicated that they had no HS diploma or GED or certificate. All others were divided into two groups: those who had received an HS diploma (HSDEG=1; there were 7,554 of these) and those remaining (862). For the first group, high school graduation year was modeled as a function of age using simple linear regression. The model was

GradYr= 2017.305787-1.001766*Age
Graduation year was rounded to the nearest whole year, and it was not allowed to exceed the year 2000, which was the most recent year of an existing study respondent. (There were seven study respondents, aged 15 or 16 , whose graduation years were set to the year 2000 in this manner). The modeling utilized only observations with non-missing age and HS graduation year for those which had indicated $\operatorname{HSDEG}=1(49,673)$. The $\mathrm{R}^{2}$ was 0.994559 , and the MSE was 0.423 . The earliest year set in this way was 1931 for two study respondents, both 86 years of age.
The remaining missing values were imputed using weighted hot deck procedures.
Imputation classes: A cross-classification of
Type of high school degree
Age (categorical)

## Sorting Variable:

Age
(17)

| Variable Name: | INDEPINC |
| :--- | :--- |
| Description: | Student s income |
| Data Used in Imputations: | Study respondents, independents (35,600) |
| Number Missing: | $8,761(24.61 \%)$ |
| Imputation Classes: | 54 CHAID segments defined by cross-classifications of |
| CPS record indicator |  |
| Pell grant status |  |
| Dependents indicator |  |
| Stafford loan status |  |
| Student marital status |  |
| Age (categorical) |  |
| Local residence |  |
| Institution level of instruction |  |
| Hispanic indicator |  |
| Institution control |  |
| White race indicator |  |
| Region |  |
| Attendance status |  |
| Asian race indicator |  |
| Citizenship |  |
| Gender |  |
| Student type |  |
| Sorting Variables: |  |
| Attendance status |  |
| Age |  |

(18) Variable Name:

Description:
Data Used in Imputations:
Number Missing:

EFC4
Expected family contribution
Study respondents $(61,767)$
29,086 (47.1 percent) total; Specifically, 10,207 (39.5 percent)
dependents, 10,743 ( 55.0 percent) independents without dependents, and 8,136 (49.5 percent) independents with dependents.

## Imputation procedure:

Records with a recorded value (EFC1) were divided into the three categories of EFC formula types, and separate regression equations were developed.

For both types of independent students the variables used for the estimation were:

Student total income
Student marital status
Student family size
Student number in college
Dummy variable for total income of $\$ 75,000$ or more

[^4]For dependent students the variables were:

Parent total income
Parent income squared
Parent family size
Parent number in college
Parent marital status
Dummy variable for total income of $\$ 75,000$ or more

Logistic regression was used to predict whether or not the student fell into the zero EFC group. If the estimated probability was below .5 , the case was estimated to have a non-zero EFC. If the value was greater than or equal to .5 , the case was estimated to have a zero EFC . For the non-zero cases, an OLS based regression formula was used to estimate the EFC. The adjusted R squared values for the OLS regressions were .69 for dependent students, .59 for independent students without dependents, and .60 for independent students with dependents.

The correlation coefficients between estimated and actual EFC were:

Dependent ..... 85

Independent/no dependents .72
Independent/with dependents . 78
For independent students, about 70 percent of the predicted values were within one thousand dollars of the actual value for the EFC. The results for dependent students were less satisfactory, with only about $28 \%$ of the values within one thousand dollars.
Table K－1．－Distribution of imputed variables before and after imputation

|  |  |  | － | \％ | $\infty \times$ | $\wedge$ | ミのo\＃\＃m | ¢～－ | 8 \％ | n ลิ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | $\begin{aligned} & i n \underset{\sim}{i} \underset{\sim}{\infty} \underset{\sim}{\infty} \\ & \underset{\sim}{n} \\ & \underset{\sim}{n} \end{aligned}$ | $\begin{aligned} & 0 \\ & \text { in } \\ & \text { in } \end{aligned}$ | $\begin{gathered} \text { oos } \\ \underset{\sim}{\infty} \underset{\sim}{n} \\ \hline \end{gathered}$ | $\stackrel{\circ}{\infty}$ | $\begin{aligned} & \circ \text { on on } \\ & \text { ond } \\ & \text { on } \\ & =1 \end{aligned}$ | $\begin{aligned} & \circ 80 \\ & \text { Nin } \\ & \text { in } \end{aligned}$ | $\stackrel{8}{\circ} \mathrm{O}$ | $\stackrel{\circ}{\sim}$ |
|  |  |  | 문 꾼 $\circ$ ○ へ 入 | $\begin{gathered} \underset{\sim}{\dot{\sim}} \stackrel{\rightharpoonup}{n} \\ \underset{\sim}{n} \end{gathered}$ | $\stackrel{\sim}{\infty} \underset{\sim}{\sim}$ | $\underset{\substack{\infty \\ \underset{\sim}{\infty} \\ \underset{i}{\infty} \\ \hline}}{ }$ |  |  |  | $\stackrel{ \pm}{\sim}$ |
|  |  |  |  | $\begin{gathered} \hat{o} \\ \text { in } \\ \text { Î } \end{gathered}$ | Co | $\stackrel{N}{\hat{A}}$ | $\underset{\infty}{\text { N}} \underset{\sim}{\text { N}} \underset{\sim}{N} \underset{\sim}{n} \text { in }$ | $\begin{aligned} & \infty \\ & \stackrel{\infty}{大} \text { o } \\ & \text { in } \end{aligned}$ |  | $\cdots$ |
|  |  | $\left\lvert\, \begin{gathered} 0_{0}^{0} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{gathered}\right.$ | べッニコ | $\ddagger$ | $\min ^{n}$ | $\underset{\infty}{\infty}$ | ミッ | 척 | $\cdots \stackrel{ }{N}$ | ช $\stackrel{\text { ¢ }}{ }$ |
|  |  |  |  | $\begin{aligned} & \stackrel{\circ}{\circ} \\ & \stackrel{i}{\hat{N}} \end{aligned}$ |  | $\begin{aligned} & \circ 8 \\ & \text { ig } \\ & \text { in } \end{aligned}$ |  | $\begin{aligned} & 8 \text { Bo } \\ & \text { Bi } \\ & \text { mis } \end{aligned}$ | $\begin{aligned} & \stackrel{\circ}{y} \\ & \underset{\sim}{n} \\ & \underset{\sim}{n} \end{aligned}$ | O－ $\cdots$ $\cdots$ $\sim$ |
|  |  |  | $\begin{aligned} & \text { on } \underset{\sim}{\infty} \underset{\sim}{\infty} \underset{\sim}{=} \underset{\sim}{=} \end{aligned}$ | $\begin{aligned} & \underset{\sim}{\underset{\sim}{\sim}} \\ & \underset{\sim}{\sim} \end{aligned}$ | $\stackrel{\leftrightarrow}{\infty} \underset{\sim}{\infty} \underset{\sim}{\infty} \underset{\sim}{\underset{\sim}{c}}$ | $\stackrel{\sim}{\underset{\sim}{i}} \underset{\stackrel{\infty}{\infty}}{\stackrel{\circ}{\infty}}$ |  |  | $\underset{\sim}{n} \underset{\sim}{\stackrel{\rightharpoonup}{\mathrm{C}}}$ | $\begin{array}{cc} \substack{0 \\ 0 \\ \dot{c} \\ \dot{\sim} \\ \\ \hline} \end{array}$ |
|  |  |  |  | $\begin{aligned} & \pm i n \\ & \sim \\ & \underset{\sim}{n} \stackrel{n}{\infty} \end{aligned}$ | $\begin{aligned} & \hat{\alpha} \approx \approx \\ & \dot{F} \underset{\sim}{2}= \end{aligned}$ | $\begin{gathered} \underset{\sim}{c} \\ \stackrel{\rightharpoonup}{7} \end{gathered}$ |  |  |  |  |
|  |  |  |  | Fin | たnm | $\simeq \infty$ | $\hat{N}$ | ホ～ | $\underset{\sim}{\sim}$ | \％ |
|  |  |  |  | $\begin{aligned} & \text { or g } \\ & \text { ond } \\ & \text { cin min } \end{aligned}$ | 웅 역 ஸi N |  |  |  | 成 | 民， |
|  |  |  |  |  |  | $\stackrel{8}{\stackrel{8}{i}} \stackrel{+}{\infty}$ |  |  | $\begin{gathered} \stackrel{\rightharpoonup}{\infty} \\ \underset{N}{\mathrm{~N}} \end{gathered}$ |  |
|  |  |  |  | $\begin{aligned} & \bar{i} \\ & \vec{i} \\ & \underset{i}{n} \end{aligned}$ | $\begin{aligned} & \text { anc } \\ & \text { in } \\ & \text { むin } \end{aligned}$ |  |  |  | $\begin{aligned} & \omega_{0} \\ & \cdots \\ & \infty \\ & \infty \end{aligned}$ |  |
|  |  |  |  |  |  |  |  |  |  |  |

Table K－1．－Distribution of imputed variables before and after imputation－Continued

|  |  |  | のパ¢ | ら | すーNの\＃ | $\wedge{ }_{\infty}$ | ¢¢ ${ }^{\text {¢ }}$ | $\stackrel{\text { N }}{\sim}$ | $\infty$ min |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | 扁名㣻 | $\begin{aligned} & \text { BO } \\ & \text { ñ } \\ & \text { nin } \end{aligned}$ |  | Oio | $\underset{\rightarrow}{\circ} \underset{\rightarrow}{\circ} \underset{\sim}{\circ} 8$ | $\stackrel{8}{8}$ | 서억잉ㅇㅡㅏ |
|  |  |  |  | $\begin{gathered} o \\ \underset{\sim}{\sim} \\ \underset{\sim}{N} \\ \underset{\sim}{c} \\ \underset{\sim}{\infty} \\ \hline \end{gathered}$ |  | $\underset{\sim}{\underset{\sim}{\infty}} \underset{\sim}{\infty} \stackrel{n}{n}$ |  | $\stackrel{\sim}{\underset{\sim}{\sim}} \stackrel{\infty}{\stackrel{~}{\sim}}$ | $\underset{\sim}{\sim} \sim_{n}^{n} \stackrel{\infty}{\sim} \underset{\sim}{c}$ $\uparrow \underset{m}{\infty} \underset{\sim}{\circ} \dot{\sim}$ |
|  |  | $\begin{aligned} & \stackrel{8}{5} \\ & 0 \\ & 0 \\ & 0 \\ & 0 \\ & 0 \end{aligned}$ | $\stackrel{N}{n}$ |  | $\underset{\infty}{\infty} \circ \dot{\sim}$ | 첟 |  | ลิठ |  |
|  |  |  | ヲ̇へ | シำ | ぶn\＃－ | $\bigcirc \bigcirc$ | タñm | $\stackrel{\sim}{\sim}$ | $\infty \times \underset{\sim}{\sim} \sim^{\infty} \times$ n |
|  |  |  |  |  |  |  |  | $\begin{aligned} & 80 \\ & 0_{0} \\ & =6 \end{aligned}$ |  |
|  |  | $\left\lvert\, \begin{gathered} 0 \\ 0 \\ 0 \end{gathered}\right.$ | $\begin{aligned} & \text { on } \stackrel{\rightharpoonup}{\infty} \\ & \stackrel{\sim}{\sim} \\ & \stackrel{\infty}{c} \end{aligned}$ |  |  |  |  |  | $\because \underset{\sim}{\sim} \sim_{n}^{\infty} \infty$ <br>  |
|  |  | $\begin{aligned} & \stackrel{8}{0} \\ & 0 \\ & 0 \\ & 0 \\ & \text { En } \end{aligned}$ |  |  |  |  |  |  |  |
|  |  |  | ๆ ${ }_{\text {® }}^{\sim}$ | กニコ | ぶの\＃N－ | に ${ }^{\text {a }}$ | ¢ ${ }_{\text {d }}$ | $\underset{\sim}{\sim}$ | ¢ ${ }_{\text {N }}$ |
|  |  | $\begin{aligned} & \stackrel{8}{n} \\ & \text { N } \\ & \text { in } \\ & \text { in } \end{aligned}$ |  |  |  |  |  | $\begin{array}{ll} \circ & 0 \\ n_{2} \\ \alpha_{2} \end{array}$ |  |
|  |  |  | $\stackrel{\circ}{\circ} \stackrel{\circ}{\circ}$ |  | $\underset{\sim}{n} \underset{\sim}{\circ} \underset{\sim}{\circ} \underset{\sim}{\circ}$ |  |  | $\stackrel{\rightharpoonup}{\mathrm{N}} \stackrel{\circ}{\circ}$ |  |
|  |  |  | $\underset{\sim}{2} \underset{i}{n}$ $\underset{\sim}{n} \underset{\sim}{n}$ |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |

Table K－1．－Distribution of imputed variables before and after imputation－Continued

|  |  |  | $\wedge$ のの®ペ |  | ミロッニの | ๑๐ヲスヘフニ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  | 스상응 |  |  |  |
|  |  |  |  |  －0 o－i i |  |  |
|  |  |  | ぶャスちion |  |  |  |
|  |  |  | のヌへニコロ | $\pm$ コニロえへに |  | 윽＝－${ }_{\text {a }}$ |
|  |  |  |  |  |  |  |
|  |  |  | ボ す す $\infty$ <br>  | Nan <br> $\dot{\sim} \dot{\sim}=0 \dot{\sim} \dot{\sim}$ |  ベ～$\dot{\sim}-\dot{\sim}$ |  <br>  |
|  |  |  |  |  |  |  |
|  |  |  | ヘラベニニセ |  | のロッへのへ | $コ の=ワ$ このa |
|  |  |  |  |  |  |  |
|  |  | $\left\|\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{array}\right\|$ |  | $\pm \infty{ }_{\infty}^{\infty} \sim \pm \underset{m}{\infty}$ $\underset{\sim}{\dot{j}} \dot{\sim} \dot{\sim} \underset{\sim}{\infty} \underset{\sim}{\infty}$ | $\bar{\infty}_{\infty}^{\infty} \stackrel{\infty}{\stackrel{\infty}{\infty}} \underset{\infty}{\infty}$ <br>  |  $\underset{\sim}{\infty} \dot{\sim}$ |
|  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |

[^5]
[^0]:    ${ }^{1}$ A student was classified as either an undergraduate, graduate, or first-professional student
    ${ }^{2}$ Study respondents were placed in one of five categories, one category for each of the five races. Whenever a study respondent's response was some multiple configuration of races, the most "minority" race (the one race within the configuration with the fewest respondents) was assigned. This hierarchy, from most "minority" to least, was American Indian, Pacific Islander, Asian, black, and white.

[^1]:    ${ }^{3}$ Alaska and Hawaii were placed in the region for outlying areas, along with Puerto Rico

[^2]:    ${ }^{4}$ Three broad age categories were defined: 1) less than or equal to 23 years; 2) between 24 and 29 years, inclusive; and 3) greater than 29 years of age

[^3]:    ${ }^{5}$ Study respondents were placed in one of five categories, one category for each of the five races. Whenever a study respondent's response was some multiple configuration of races, the most "minority" race (the one race within the configuration with the fewest respondents) was assigned. This hierarchy, from most "minority" to least, was American Indian, Pacific Islander, Asian, Black, and White.

[^4]:    ${ }^{6}$ Attendance status was used as a sort variable in addition to defining CHAID segments because attendance status was an important variable for determining student's income and it was not included in all CHAID segments.

[^5]:    ${ }^{\text {\＃}}$ Less than 0.5 percent
    ${ }^{1}$ Percentages may not sum to 100 due to rounding．
    ${ }^{2}$ Graduates／first－professionals are independent by
    ${ }^{2}$ Graduates／first－professionals are independent by definition．However， 353 of them were coded as dependents before imputation and all 353 had missing parents＇income． NOTE：To protect confidentiality，some numbers have been rounded．Following data imputations，variables were reviewed and revised，if necessary，to adjust for

    SOURCE：U．S．Department of Education，National Center for Education Statistics，National Postsecondary Student Aid Study，1999－2000（NPSAS：2000）

