

INVESTIGATING NONRESPONSE BIAS: AN ANALYSIS OF EARLY, MID- AND LATE RESPONDERS  
IN A PANEL SURVEY OF TEXAS HIGH SCHOOL STUDENTS

AAPOR ANNUAL CONFERENCE

MONTREAL, CANADA

MAY 2006

DRAFT PAPER

JULIE PAASCHE & ROBERT SANTOS

NUSTATS

## INTRODUCTION

The Texas Higher Education Opportunity Project (THEOP) is a multi-year study beginning in the fall of 2000. The study investigates educational opportunities and decision-making under a policy that guarantees admission to any Texas public college or university to high school seniors who graduate in the top ten percent of their class. One aspect of the study involves administrative data from ten colleges and universities in Texas of varying selectivity but including the two flagship Texas universities. The core of the study is a longitudinal survey of students who were sophomores and seniors enrolled in Texas public schools in the spring of 2002.

TABLE 1: TWO-COHORT LONGITUDINAL SURVEY

	2002	2003	2004	2006
<b>Senior Cohort</b>	Baseline (Wave 1)  N=13,803	Wave 2  N=5,836		Wave 3  N=~5,800
<b>Sophomore Cohort</b>	Baseline (Wave 2)  N=19,969		Wave 2  N=3,092	

The baseline surveys of seniors and sophomores (Wave 1) were conducted in spring, 2002, consisting of 13,803 seniors and 19,969 sophomores based on a stratified random sample of 105 public high schools in the state of Texas (see Table 1). The baseline surveys were mainly collected in-class using a self-administered scannable booklet. The instrument queried students about their course taking and grades, experiences with guidance counselors, knowledge and perceptions of college admissions process, college perceptions, future plans and demographic information, including race and ethnic origin, family background, and household structure. Seniors (but not sophomores) were asked a battery of questions about colleges applied to, their self-reported admission status, and plans to attend college.

Subsamples of the senior and sophomore cohorts were selected for follow-up interviews. For the senior cohort 5,836 respondents re-interviewed (Wave 2) one year after graduating from high school to ascertain primary post-secondary school activity, military enlistment, labor force participation, etc. They are currently being re-interviewed (Wave 3) during spring, 2006, when a large majority of those who attended college are juniors and seniors. For sophomores, 3,092 respondents from the sophomore cohort were re-interviewed (Wave 2) during their senior year to determine whether they remained the same school (were “stayers” or “leavers”), and to record their progress in high school, their college plans, and changes in other circumstances.

This paper focuses on the senior cohort. The survey universe for the Senior Wave 2 Study was defined as all students who participated in the baseline survey, signed the consent form, and could be identified by name. The vast majority of senior baseline participants provided at least some information on the contact information sheet included in the baseline survey. For those participants who signed the assent form but failed to provide any contact information, we assigned the home address from school directories to the extent possible and also conducted research, mainly via the

web, for contact information. Baseline respondents who were anonymous (gave no identifying information) or who failed to sign the written assent form were not included for selection.

Over 131,000 calls were made to complete 5,836 interviews. An average of fifteen call attempts were made per piece of sample. For completed surveys, the average was 12 call attempts while for those cases that never completed a survey the average is over 20 call attempts, ranging from 1 attempt to 99 attempts. The average length of completed full interviews was approximately 25 minutes (ranging from 9 minutes to 64 minutes).

Incentives were used to increase participation rates at the tail end of the field period. For military proxy surveys \$5 was sent with a thank-you letter to generate goodwill among proxies that should lead to an increased respondent participation rate in the full survey. For full surveys \$10 was sent in a thank-you letter for surveys completed from September 15 to October 16. From October 17 to December 9, a \$20 incentive was used. The \$20 incentive will also be offered to military proxy cases as we attempt to complete full surveys with respondents throughout 2004.

A total of 5,836 interviews were completed with 5,741 completed as full interviews with the respondent and 95 interviews completed with proxies for respondents that are unreachable due to military service (see Table 2). Only 159 respondents (2%) refused to participate in the study and 5 cases were ineligible. Of the remaining sample, roughly two-thirds could not be located to complete the survey while one third were believed to be 'located' but did not complete the survey during the field period.

TABLE 2: CALL OUTCOMES

Call Outcome	Frequency	Percent
Completed Full Survey	5741	69
Completed Proxy Survey	95	1
<b>TOTAL COMPLETES</b>	<b>5836</b>	<b>70</b>
<b>REFUSED TO PARTICIPATE</b>	<b>159</b>	<b>2</b>
Deceased	4	0
Special Education Student	1	0
<b>TOTAL INELIGIBLE</b>	<b>5</b>	<b>0</b>
<b>REMAINING SAMPLE</b> (1/3 Located, 2/3 Not Located)	<b>2346</b>	<b>28</b>
<b>GRAND TOTAL</b>	<b>8346</b>	<b>100</b>

Investigating the potential for non-response bias in Wave 2 can inform fielding decisions for Wave 3 of the study. Such an analysis facilitates the development of an efficient field protocols, the setting of optimal targeted response rates, and the delineation of trade-offs between data collection costs and data quality (e.g., are X thousand dollars worth a Z percent increment in the response rate given what we know about the nature of the nonresponse).

An analysis of surveys completed in the early portion of data collection, mid and late yields useful information on whether late respondents differ from their counterparts that completed the survey earlier. Interviews that are conducted late in the field period are much more expensive to obtain than their earlier counterparts. By knowing the extent that “early” respondents are similar or different than “later” respondents, informed decisions can be made about the benefit of incrementing the response rate by a percentage point or two at the end of data collection for subsequent waves of the study. Thus, the objective of this analysis is twofold -- to document the potential for nonresponse bias in the Wave 2 senior follow-up, and to guide decisions on Wave 3 data collections in terms of the value of getting the last percentage or two of response rate relative to the cost.

## SUMMARY OF RESULTS

The analyses presented here support a hypothesis of minimal or negligible nonresponse bias on a variety of key survey variables. Variables that were reviewed include demographic, educational, employment, high school experiences and self-esteem variables. The analysis was conducted by dividing the first follow-up senior data into quartiles based on the date the survey was completed with the respondent. The pattern of responses across the quartiles was then compared. The main findings are as follows:

- On all selected demographic variables except gender, the pattern of response was the same throughout data collection.
- For gender, the last quartile of respondents contained a higher percentage of male respondents (52% male in the last quartile compared to 47% male total). A review of the entire trajectory of cumulative surveys by gender (Figure 5) indicates that data collection by gender remained fairly constant. In fact, in the final 5% of cases collected excluding military proxy cases and those completed with military respondents in 2004 (n=287), 48% were male compared to 47% male total. In addition, the final Senior Wave 2 unweighted gender percentages were the same as the baseline unweighted percentages.
- Eleven educational variables were selected for review. The pattern of responses is consistent throughout data collection with only a few variables such as percent attending school, or granted admission through the Top Ten Percent law showing very slight divergence among the quartiles.
- Employment variables show a minimal increase in the percent of students working, however, this may be expected as more respondents find jobs as the amount of time following graduation from high school increases.
- Four high school experience variables show similar patterns of response throughout data collection.
- The pattern of responses to selected self-esteem questions was constant across the quartiles.
- An additional analysis of the last 10% of collected interviews on twenty-one variables shows mostly negligible or no differences between the first 90% and last 10% of cases. We do see a

few mild trends indicating slight divergence from the earlier cases; however, these are likely to be within the margin of error.

The remainder of this paper details the findings listed above and contains figures that illustrate the pattern of responses.

## DEFINITIONS

For this analysis, Senior Wave 2 surveys completed by proxy and surveys completed in 2004 with students in the military were excluded. The definitions of total, early, mid and late completes are as follows:

- **Total** completes are all Senior Wave 2 surveys completed in 2003 with respondents. Surveys completed by proxy or those completed in 2004 by respondents in the military were excluded. (n=5,741)
- **Early** completes are defined as surveys from the first quarter (25%) of respondents. These surveys were conducted from April 3, 2003 to April 26, 2003. (n=1,420)
- **Mid** completes are the second and third quartiles (50% of cases) combined. These surveys were conducted from April 27, 2003 to June 22, 2003. The second and third quartiles were combined throughout this report to enhance the ease of reviewing results. (n=2,880)
- **Late** completes represent the final 25% of completed surveys, including surveys conducted from June 23 to December 27, 2003. (n=1,441)

Key variables and a selection of other variables interspersed throughout the survey were selected for analysis. The *pattern* of the responses should be compared to the total for similarity or dissimilarity more intensively than comparing individual percentages.

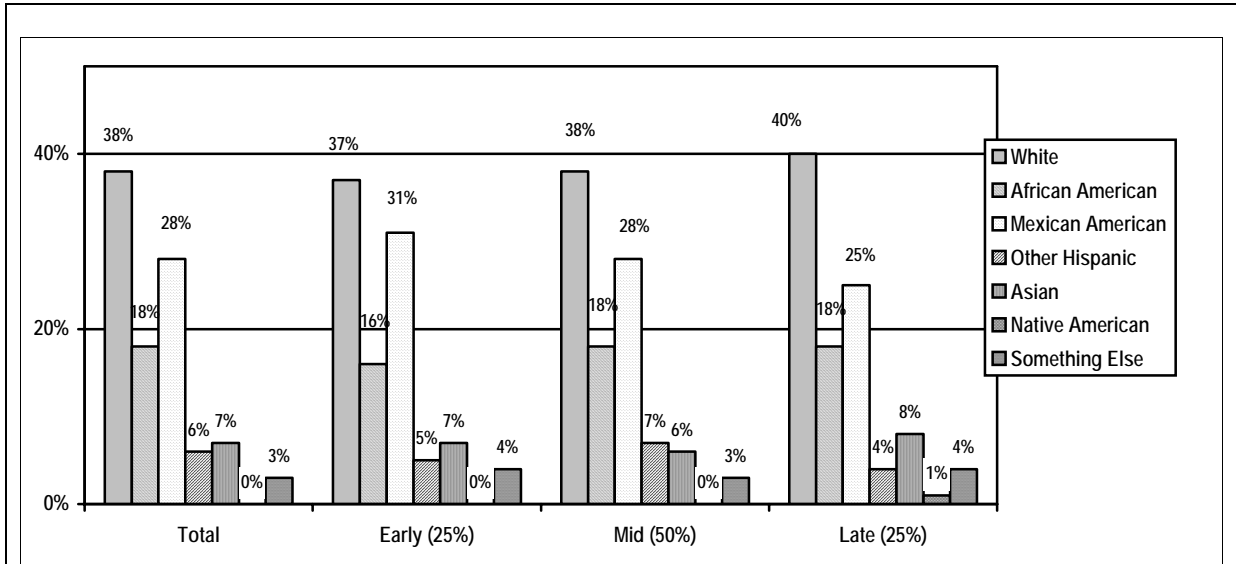
It should be noted that responses to some questions (i.e. current course taking or employment questions) might differ slightly based on the time of year the survey was completed (i.e. more students may have a summer job than are employed during the school year).

Throughout this report totals may not equal 100% due to rounding. All figures are based on unweighted data.

## DEMOGRAPHIC VARIABLES

Figure 1 shows that the pattern of race/ethnicity of respondents varies little by the date the interviews were completed.

FIGURE 1: RACE/ETHNICITY (Q38) [TOTAL N=5,741]



The Senior Wave 2 survey did not include a question on the respondent's country of birth, however it did ask for father and mother's birth countries. These variables have been used as a substitute for respondent's country of birth. In Figure 2 and 3 we see little variation in the number of foreign-born parents.

FIGURE 2: FATHER BORN IN UNITED STATES (Q39) [TOTAL N=5,741]

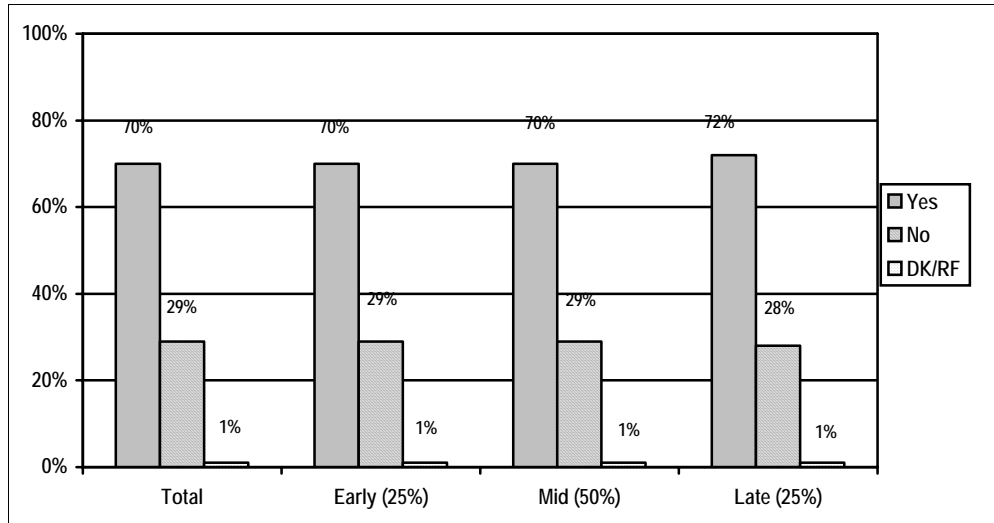
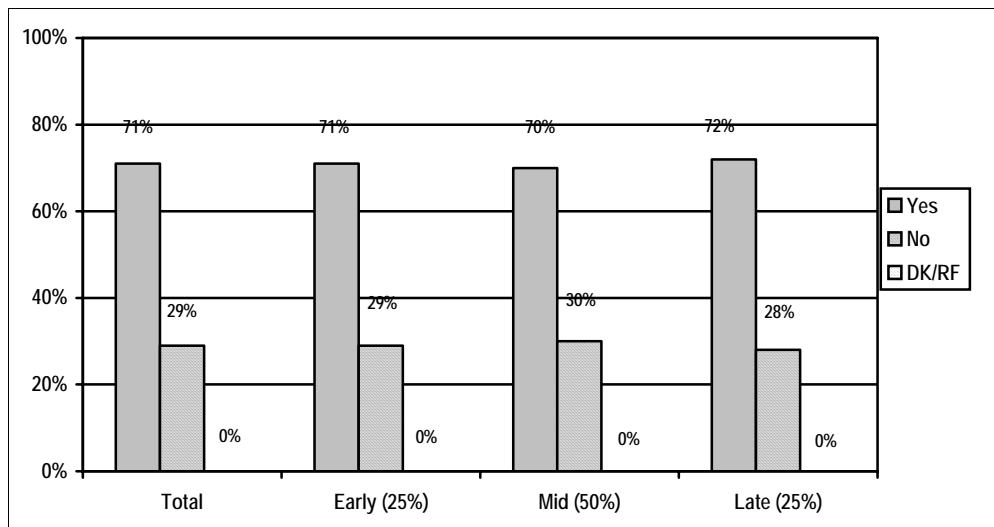


FIGURE 3: MOTHER BORN IN UNITED STATES (Q40) [TOTAL N=5,741]



Reviewing respondent's gender by date of completed survey we find a higher percentage of males completing in the last quarter of respondents (see Figure 4). For the last 5% of completed surveys the gender split is closer to the total with 48% of the final 287 surveys (excluding proxy and 2004 military cases) from male respondents and 52% with female respondents (figure not shown below).

**FIGURE 4: GENDER**

[GENDER CODED BY INTERVIEWER. CASES CODED "NOT SURE" WERE REMOVED FROM THE BASE (TOTAL N= 5,736)]

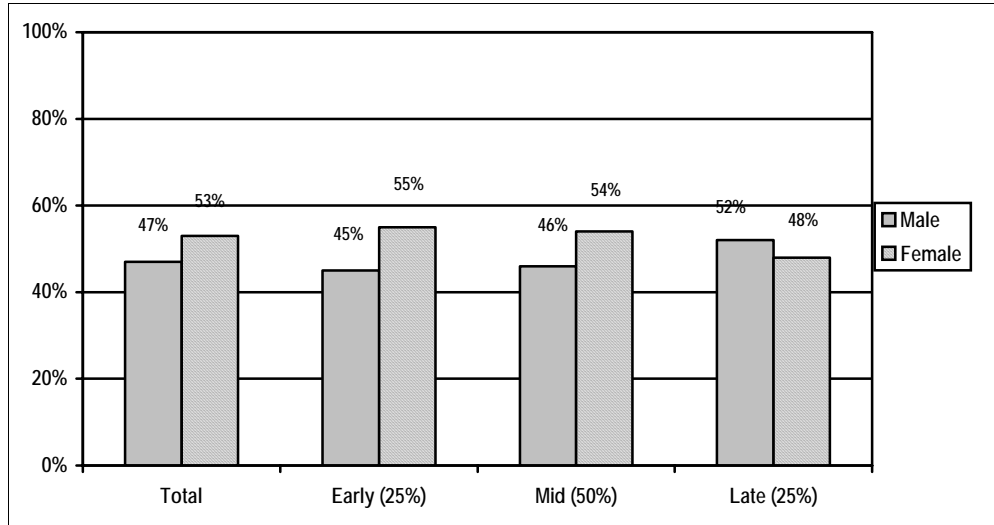
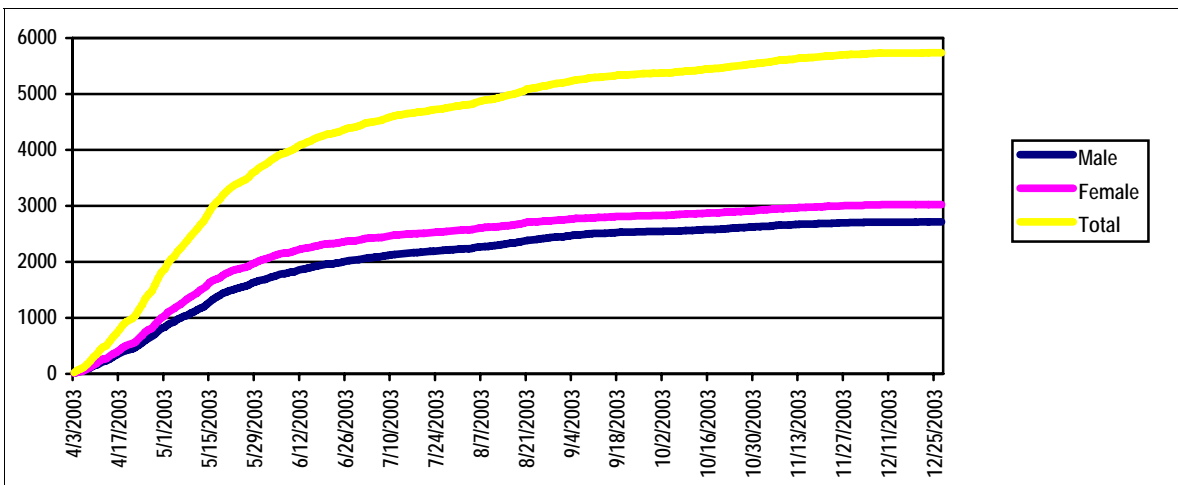


Figure 5 shows the trajectory of completed surveys by date and by gender. It is apparent that data collection by gender is consistent throughout the survey period and mirrors the baseline survey (excluding missing data from the baseline survey, 47% of baseline senior respondents were male and 53% were female). These are the same overall percentages as the Senior Wave 2 study.

**FIGURE 5: CUMULATIVE NUMBER OF COMPLETED SURVEYS BY DATE AND GENDER**

[GENDER CODED BY INTERVIEWER. CASES CODED "NOT SURE " WERE REMOVED FROM THE BASE (TOTAL N= 5,736)]





## EDUCATIONAL VARIABLES

Figure 6 demonstrates that high school graduation rates are similar for all quartiles.

**FIGURE 6: GRADUATED FROM HIGH SCHOOL (Q1) [TOTAL N=5,741]**

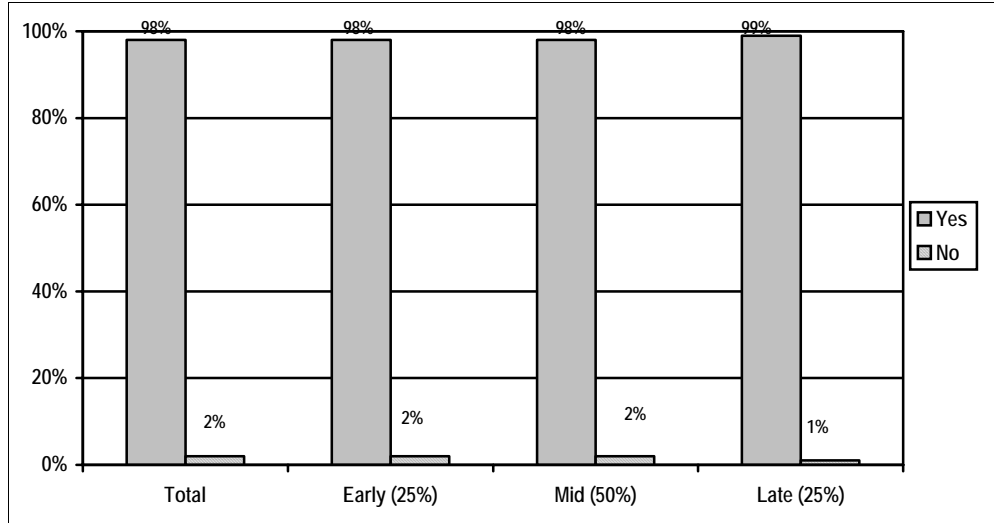
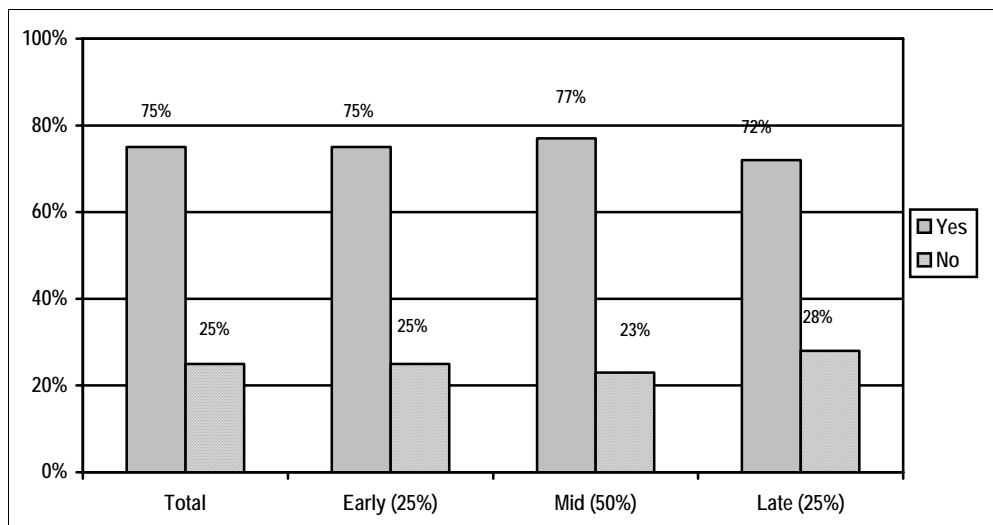


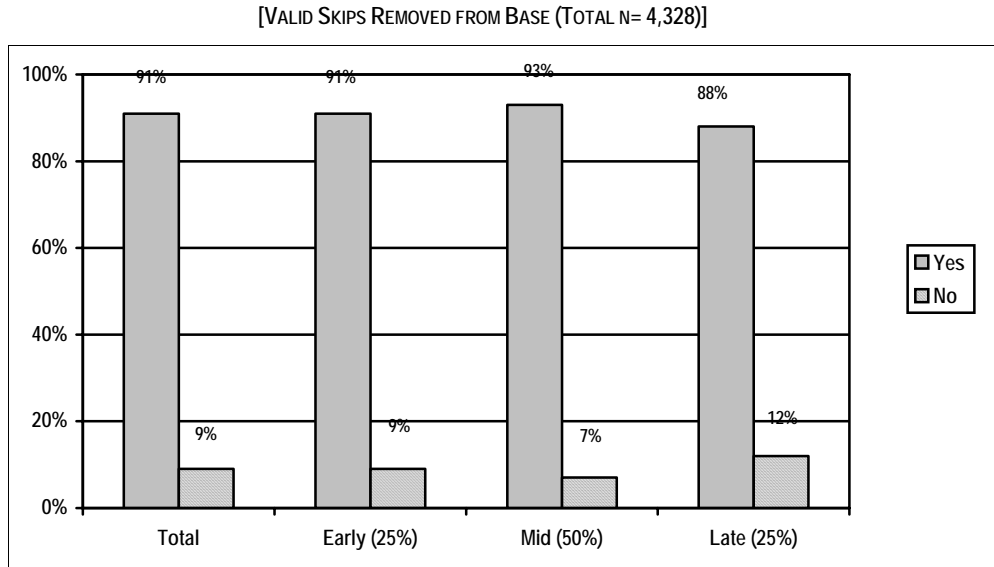
Figure 7 shows a slight decrease (72% compared to 75% total) in the percentage of respondents that have taken courses for academic credit among the last quartile of surveys collected.

**FIGURE 7: HAVE ATTENDED VOCATIONAL, TECHNICAL, OR TRADE SCHOOL, OR TAKEN COURSES FROM A UNIVERSITY OR COLLEGE FOR ACADEMIC CREDIT SINCE SEPTEMBER 2002 (Q2E) [TOTAL N=5,741]**



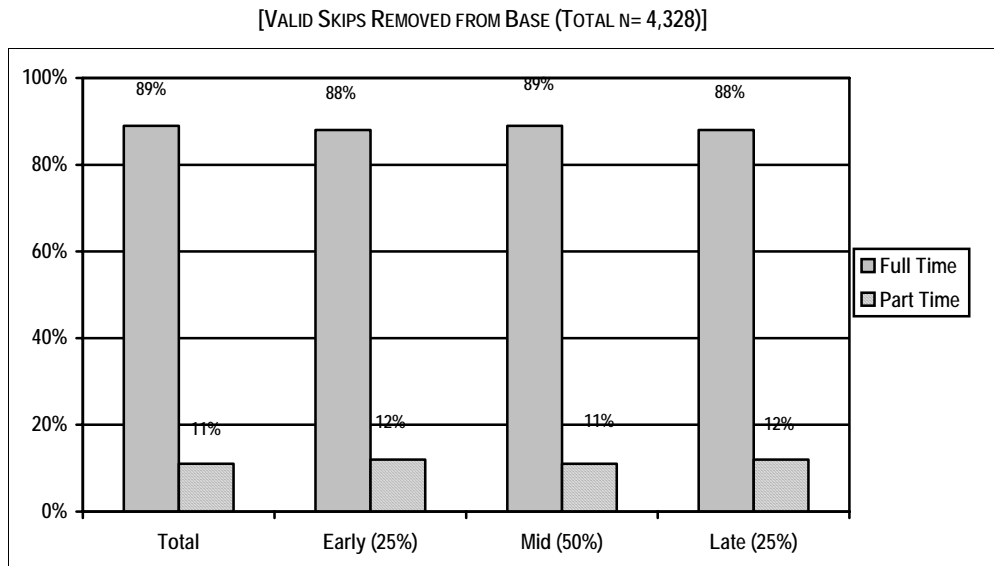
Similarly, Figure 8 shows the same pattern as the total but slightly fewer students currently attending school, however, this question (Q2e\_0. And are you currently attending school?) may be subject to the school academic calendar and the later responses were collected from June 23 to December 27.

**FIGURE 8: CURRENTLY ATTENDING SCHOOL (Q2E\_0)**



Whether students are attending school full or part time (Figure 9) is constant across data collection.

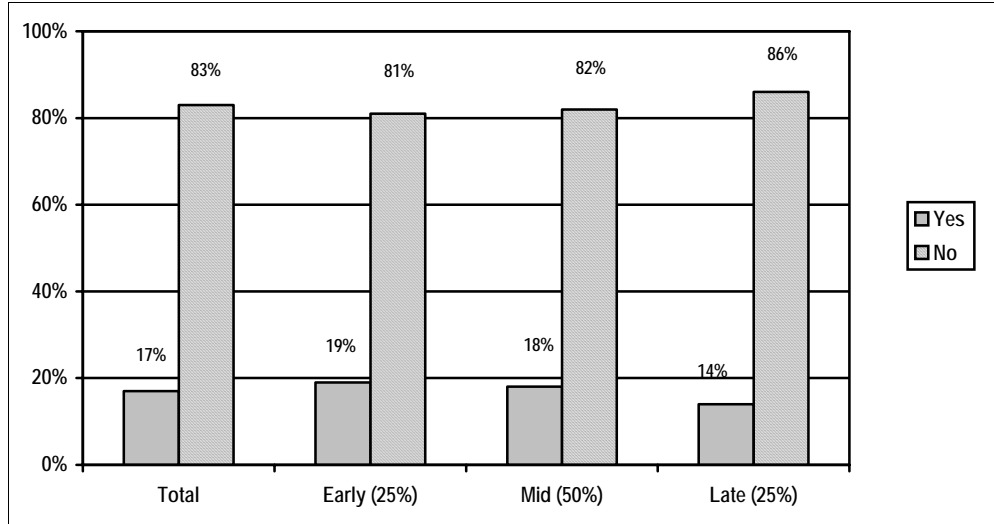
**FIGURE 9: ATTENDING FULL OR PART TIME (Q2E\_4)**



A slightly higher percentage of students that completed the survey early in data collection said they were granted admission under the Top 10 Percent Law (Figure 10) or that the Top Ten Law was a reason for selecting their school (Figure 11).

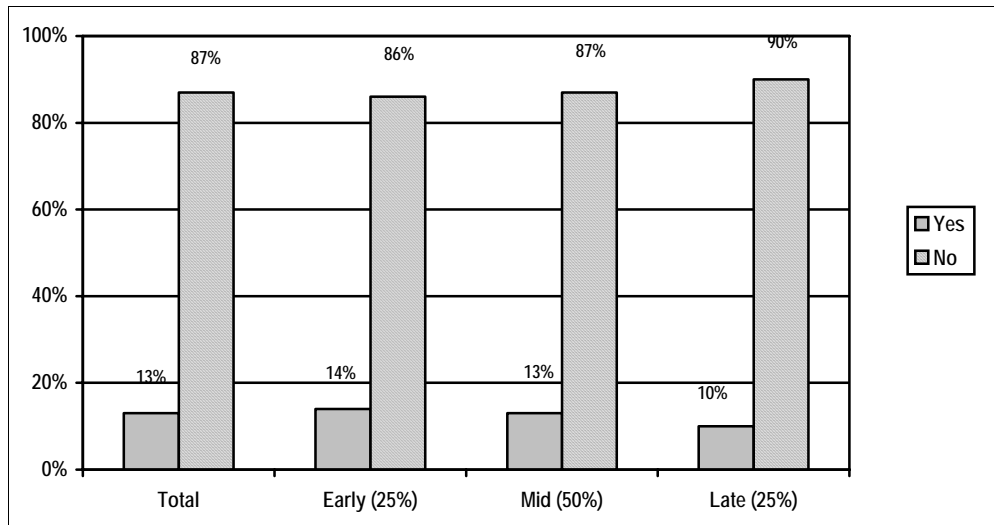
**FIGURE 10: DID ANY TEXAS COLLEGE OR UNIVERSITY GRANT YOU ADMISSION THROUGH THE TOP 10 PERCENT LAW? (Q5A)**

[VALID SKIPS REMOVED FROM BASE (TOTAL N= 4,715)]



**FIGURE 11: TOP TEN LAW WAS A REASON FOR ATTENDING <PRINCIPAL INSTITUTION> (Q6i)**

[VALID SKIPS REMOVED FROM BASE (TOTAL N= 5,143)]



Pursuit of a bachelor's or associate's degree (other degrees were not included in the figure) is quite constant (Figure 12).

**FIGURE 12: PURSUING ASSOCIATE'S OR BACHELOR'S DEGREE (Q2E\_6)**

[VALID SKIPS REMOVED FROM BASE (TOTAL N= 4,328)]

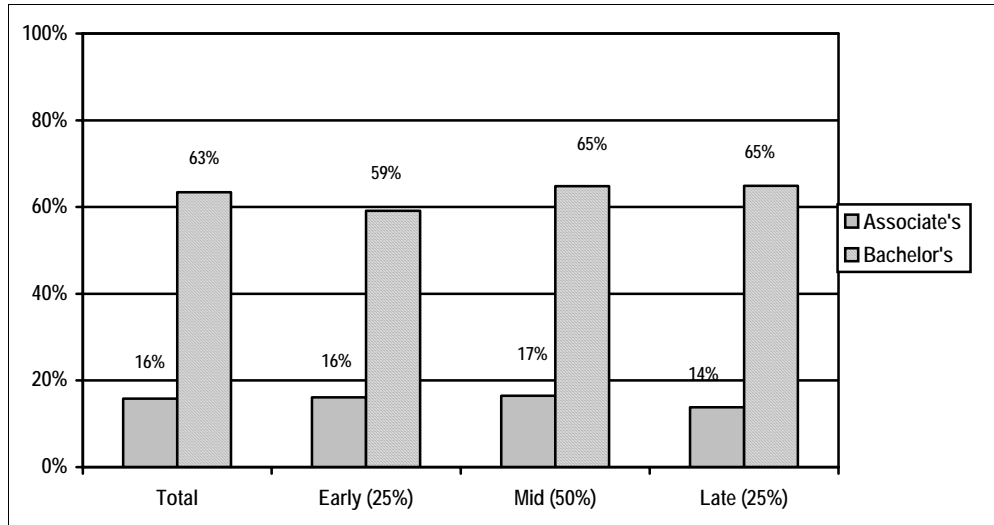


Figure 13 shows the student-reported cost to attend school including tuition, academic fees, room and board, and daily expenses for living and entertainment. The pattern is very consistent across the data collection time period.

**FIGURE 13: COST TO ATTEND SCHOOL IN ACADEMIC YEAR 2002 - 2003 (Q7)**

[VALID SKIPS REMOVED FROM BASE (TOTAL N= 4,363)]

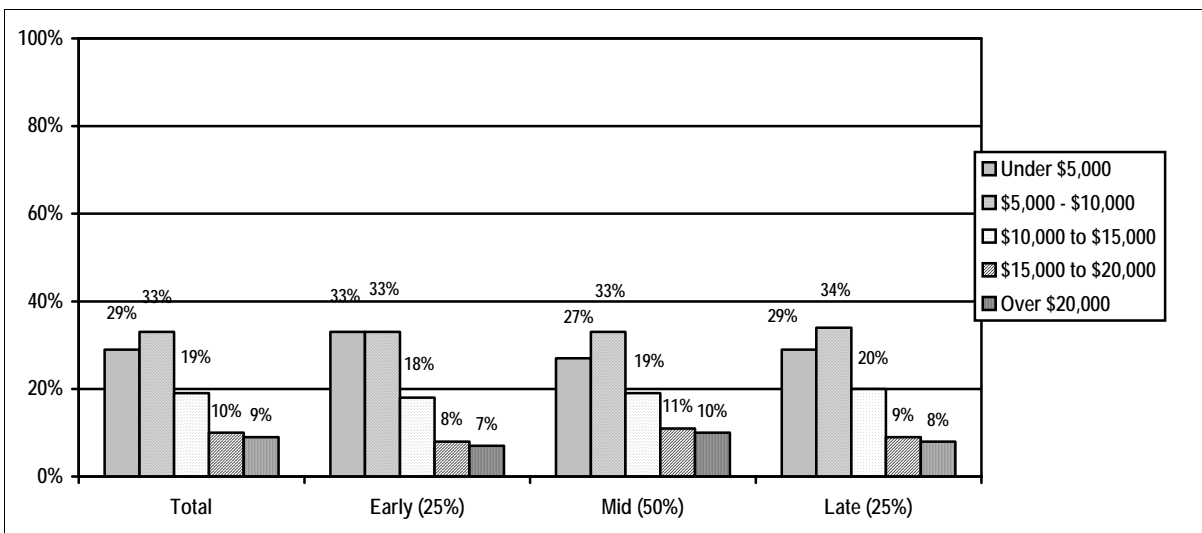
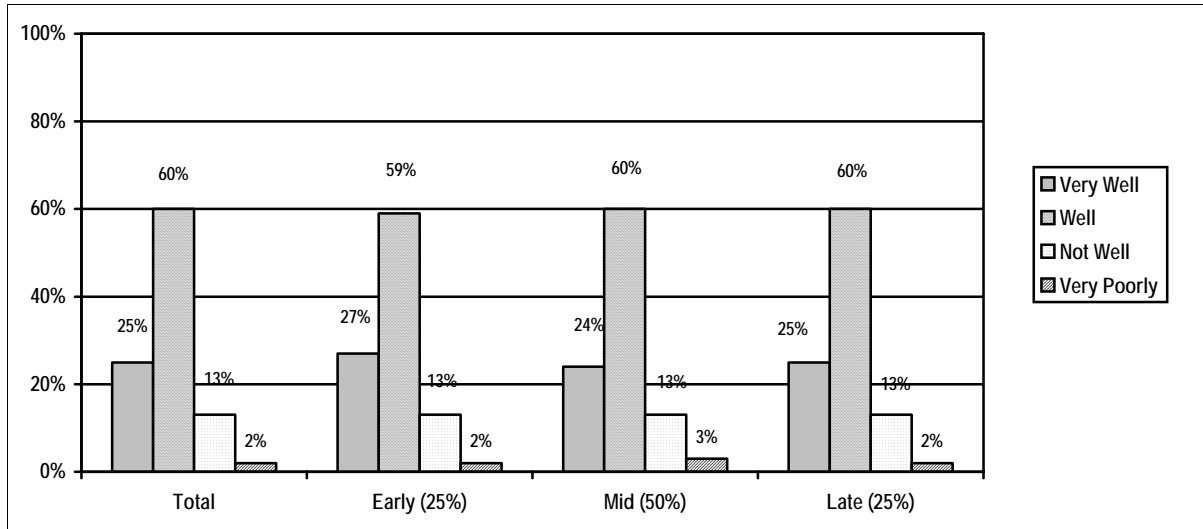


Figure 14 displays students' assessment of the preparation they received for their current academic courses. There is very little variation across the data collection period.

**FIGURE 14: PREPARATION FOR CURRENT ACADEMIC COURSES (Q16)<sup>1</sup>**

[VALID SKIPS AND MISSING RESPONSES REMOVED FROM BASE (TOTAL N= 4,361)]



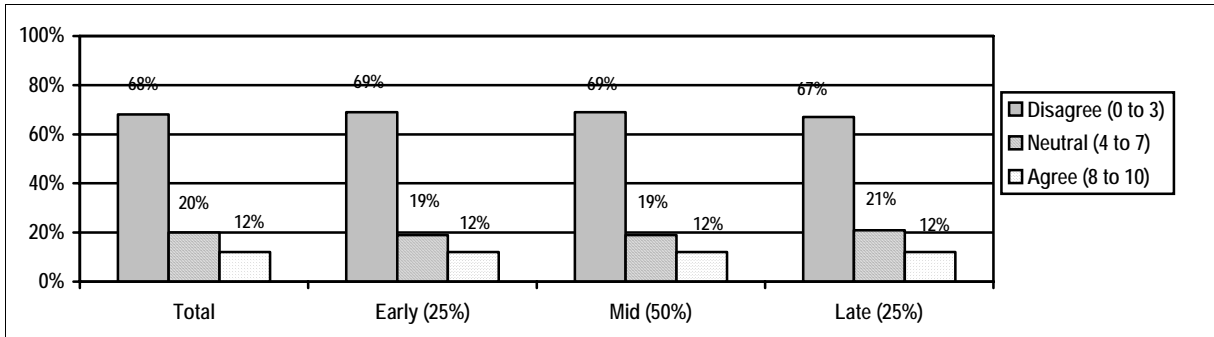
<sup>1</sup> The full text of Q16 is: "Now that you have some educational experience beyond high school, how prepared do you think you were for your current academic courses?"

Students that had attended school since the end of their senior year were given a series of statements about their school experiences and asked to indicate how much they agreed or disagreed with the statement using a 0 to 10 scale. For Figure 15 and Figure 16 the responses have been collapsed into three categories for analysis (see legend). In Figure 15 and Figure 16 we see very similar patterns of responses across data collection time periods.

**FIGURE 15: WISH HAD SELECTED ANOTHER COLLEGE**

0=TOTAL DISAGREEMENT AND 10=TOTAL AGREEMENT (Q21C)<sup>2</sup>

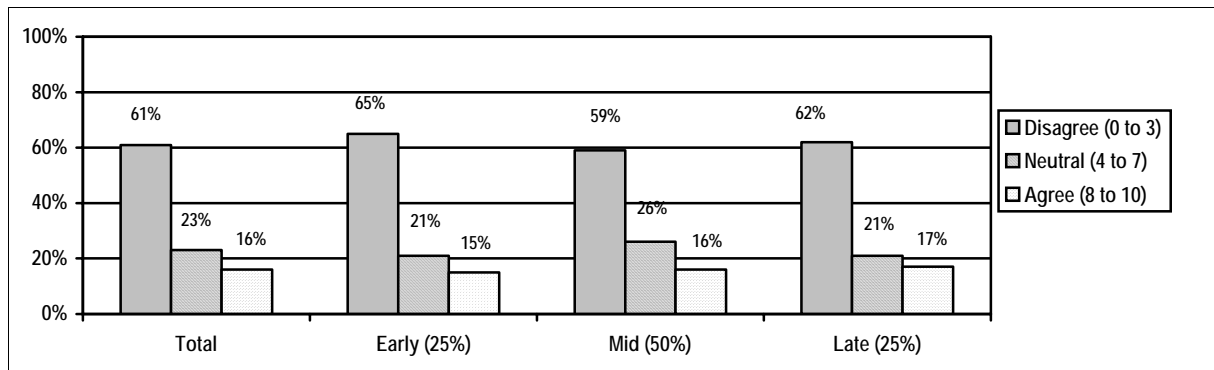
[VALID SKIPS AND MISSING RE SPONSES REMOVED FROM BASE (TOTAL N= 4,361)]



**FIGURE 16: DIFFICULT TO GRADUATE IN FOUR YEARS**

0=TOTAL DISAGREEMENT AND 10=TOTAL AGREEMENT (Q21F)<sup>3</sup>

[VALID SKIPS AND MISSING RE SPONSES REMOVED FROM BASE (TOTAL N= 4,361)]



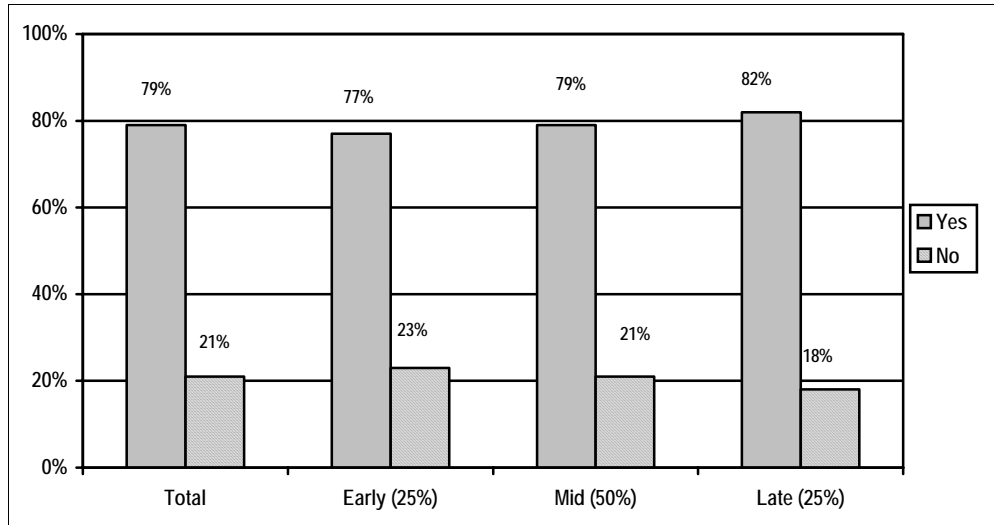
<sup>2</sup> The full text of Q21c is: “On a scale of 0 to 10, where 0 indicates total disagreement and 10 indicates total agreement, how much do you agree or disagree with each of the following statements about your experiences at <PRINCIPAL INSTITUTION>? I wish I had selected another college.”

<sup>3</sup> The full text of Q21f is: “On a scale of 0 to 10, where 0 indicates total disagreement and 10 indicates total agreement, how much do you agree or disagree with each of the following statements about your experiences at <PRINCIPAL INSTITUTION>? It will be difficult to graduate in four years.”

## EMPLOYMENT VARIABLES

Employment variables are likely to be more subject to change based on the date of data collection; however in Figure 17 we see only slight variation in the quartiles.

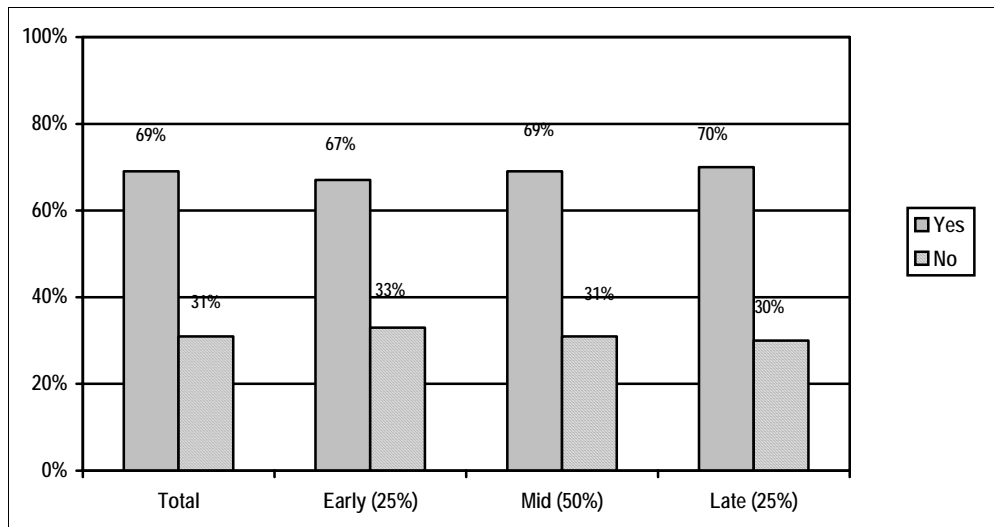
FIGURE 17: WORK FOR PAY SINCE THE END OF SENIOR YEAR (Q2B) [TOTAL N=5,741]



Again, in the question asking whether the respondent is currently working for pay, one might expect to see an increase based on the date the survey was conducted, but the variation is negligible in Figure 18.

FIGURE 18: CURRENTLY WORKING FOR PAY (Q2B\_1)

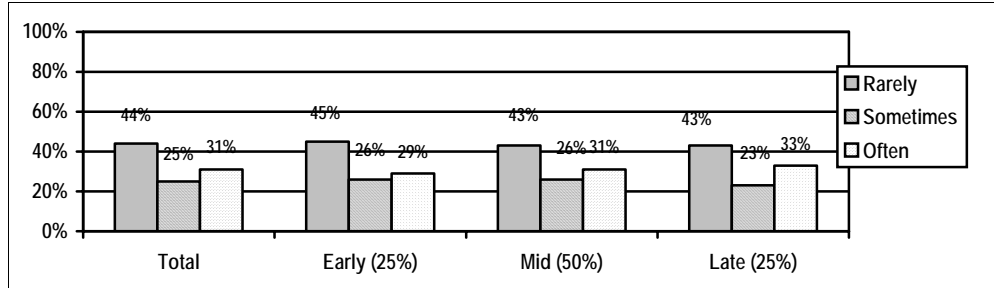
[VALID SKIPS REMOVED FROM BASE (TOTAL N= 4,546)]



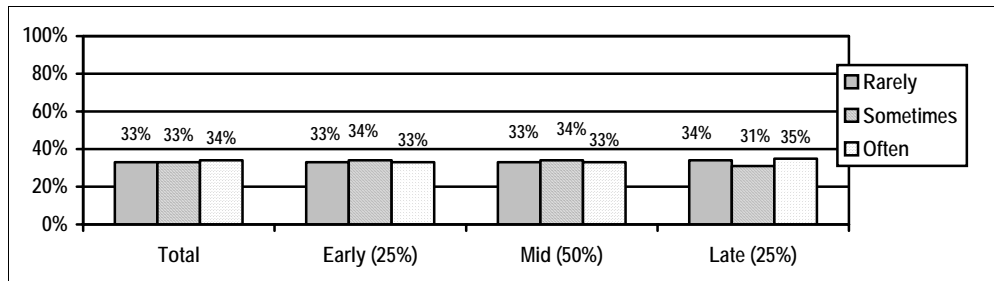
## HIGH SCHOOL EXPERIENCE VARIABLES

Figures 19 through 22 explore a few of the variables related to high school experiences. In each of these figures the patterns of responses remain constant over the data collection period.

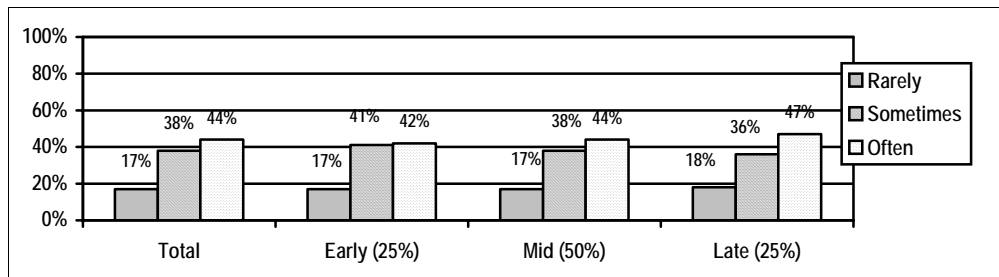
**FIGURE 19: PARENTS CHECKED IF HOMEWORK WAS DONE (Q27A) [TOTAL N=5,741]**



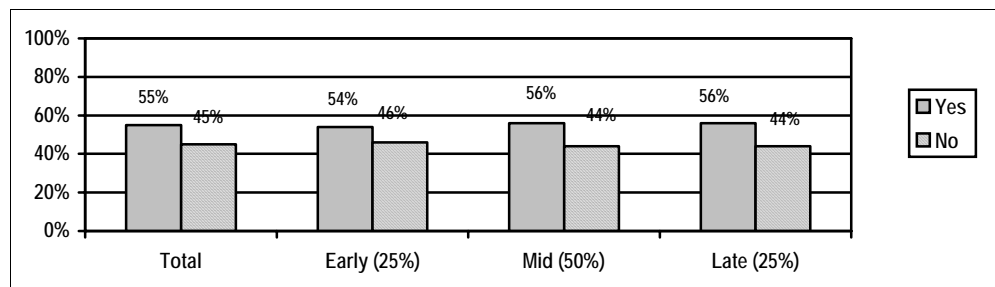
**FIGURE 20: PARENTS REWARDED FOR GOOD GRADES (Q27D) [TOTAL N=5,741]**



**FIGURE 21: MY FAMILY DID FUN THINGS TOGETHER (Q28F) [TOTAL N=5,741]**



**FIGURE 22: HAD CURFEW FOR SCHOOL NIGHTS (Q29) [TOTAL N=5,741]**





## SELF-ESTEEM VARIABLES

On a few self-esteem variables (Figure 23 and 24) the pattern of responses is similar among the quartiles.

FIGURE 23: FEEL GOOD ABOUT SELF (Q33A)<sup>4</sup> [TOTAL N=5,741]

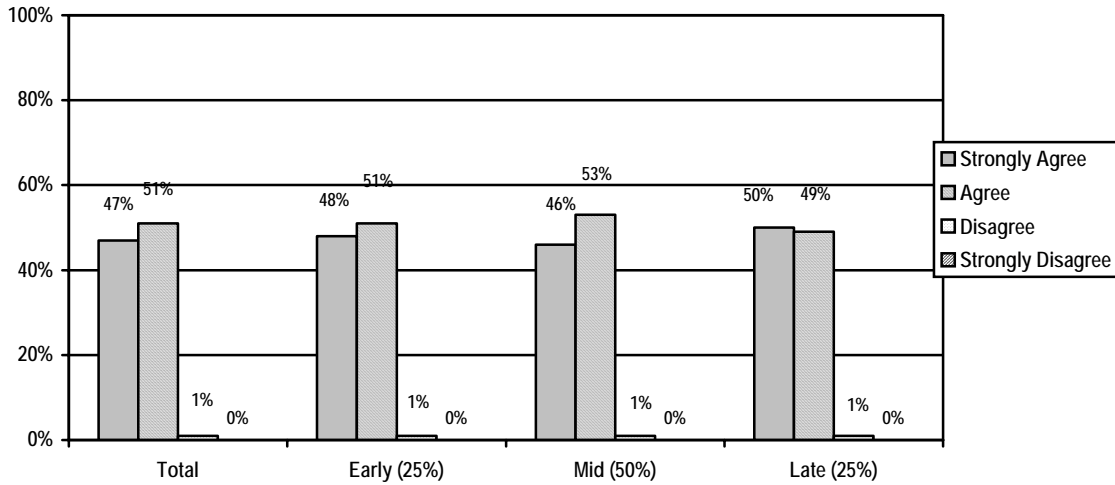
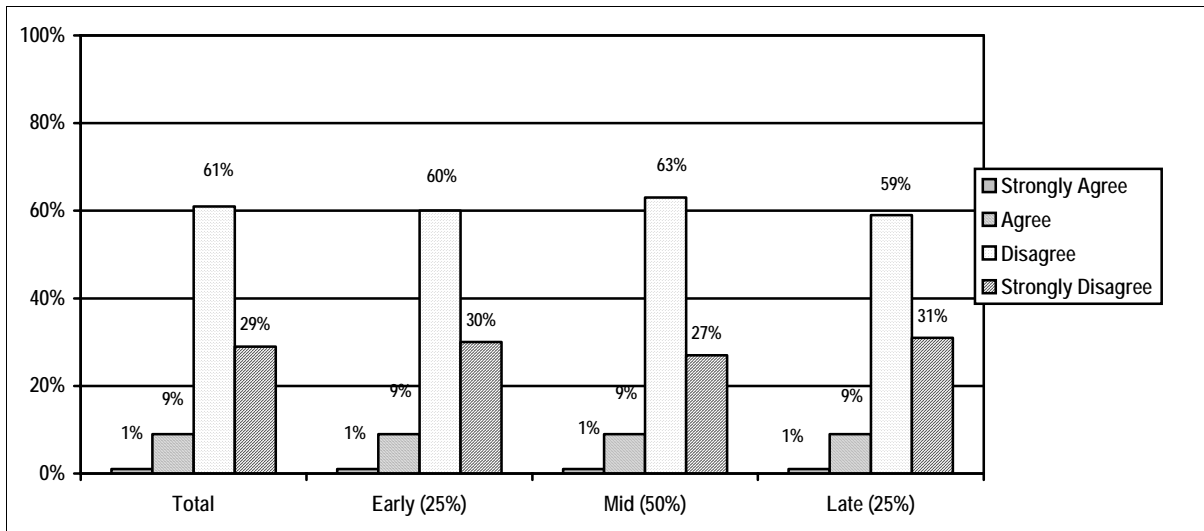


FIGURE 24: FEEL DO NOT HAVE MUCH TO BE PROUD OF (Q33i)<sup>5</sup> [TOTAL N=5,741]



<sup>4</sup> The full text of Q33a is: “For each of the following statements, please tell me whether you Strongly Agree, Agree, Disagree or Strongly Disagree. I feel good about myself.”

<sup>5</sup> The full text of Q33i is: “For each of the following statements, please tell me whether you Strongly Agree, Agree, Disagree or Strongly Disagree. I feel I do not have much to be proud of.”

## LAST 10% OF CASES COLLECTED

As an additional test of differences among early and late cases, the last 10% of cases collected were compared to the total and to the first 90% of cases. Surveys conducted with proxies for students unreachable due to military service and surveys completed with military respondents in 2004 were excluded from this analysis. The last 10% of cases (n=573) were collected from August 29 through December 27, 2003 while the first 90% of cases (n=5168) were collected from April 3 to August 28, 2003.

It appears that the last 10% of cases collected portray several trends. They are slightly more likely to be white, male, and have parents that were born in the United States. In addition, they are slightly less likely than the cases collected earlier to be attending school currently or to have been granted admission under the Top 10 Percent Law. These differences, however, are probably within the margin of sampling error. Column 5 of Table 1 presents the bias when only the first 90% of cases collected are used to estimate the population percentage. Column 6 shows the relative bias (as a percentage of the sample estimate, i.e., the ratio of the Col. 5 entry to the Col. 2 entry). The bias calculations were made for those few variables that showed at least a 5 percentage point difference between the estimates using the first 90% of cases and the last 10% of cases. The nominal bias in all cases is less than one percentage point (Col. 5), and the relative biases are all within 3 percentage points. Taken collectively (i.e., given all the other variables with virtually no impact on bias), these results suggest that the risk of faulty inference due to nonresponse bias (had we ceased data collection at 90% of the case count) would have been negligible.

TABLE 1: LAST 10% OF CASES COMPARED TO TOTAL AND FIRST 90% [TOTAL N=5,741]

Column 1	2	3	4	5	6
	Total	First 90% of cases	Last 10% of cases	Bias using first 90% <sup>6</sup>	Bias as % of estimate <sup>7</sup>
Date of interview	4/3 to 12/27 Total n=5,741	4/3 to 8/28 n=5,168	8/29 to 12/27 n=573		
<b>Variable</b>					
Race/Ethnicity – White	38%	38%	43%	-0.5%	-1.3%
Race/Ethnicity - African American	18%	18%	19%		
Race/Ethnicity - Mexican American	28%	28%	26%		
Race/Ethnicity - Other Hispanic	6%	6%	4%		
Race/Ethnicity - Asian	7%	7%	6%		
Race/Ethnicity - Native American	0%	0%	0%		
Race/Ethnicity - Something Else	3%	4%	2%		
Gender - Male	47%	47%	49%		
Gender - Female	53%	53%	51%		
Father born in United States (q39=yes)	70%	70%	77%	-0.7%	-1.0%

<sup>6</sup> “Bias Using First 90%” was calculated if the difference between the result for the last 10% of cases and the result for the first 90% of cases is 5 percentage points or greater.

<sup>7</sup> “Bias as % of Estimate” was calculated if the difference between the result for the last 10% of cases and the result for the first 90% of cases is 5 percentage points or greater.

	<b>Total</b>	<b>First 90% of cases</b>	<b>Last 10% of cases</b>	<b>Bias using first 90%</b>	<b>Bias as % of estimate</b>
Mother born in United States (q40=yes)	71%	70%	76%	-0.6%	-0.8%
Graduated from high school (q1=yes)	98%	98%	99%		
Attended school since graduating from high school (q2e=yes)	75%	76%	73%		
Currently attending school (q2e_0=yes, valid skips excluded from base)	91%	92%	83%	0.4%	0.4%
Attending school full time (q2e_4=full time, valid skips excluded from base)	89%	89%	86%		
Granted admission though Top 10 Percent Law (q5a=1, valid skips excluded from base)	17%	18%	13%	0.5%	2.7%
Decided to attend Principal Institution because of Top Ten Law (q6i=yes, valid skips excluded from base)	13%	13%	11%		
Pursuing Bachelor's degree (q2e_6=3, valid skips excluded from base)	63%	64%	61%		
Pursuing Associate's degree (q2e_6=2, valid skips excluded from base)	16%	16%	15%		
School year costs under \$5,000 (q7=1, valid skips excluded from base)	29%	29%	31%		
School year costs \$5,000 to \$10,000 (q7=2, valid skips excluded from base)	33%	33%	32%		
School year costs \$10,000 to \$15,000 (q7=3, valid skips excluded from base)	19%	19%	20%		
School year costs \$15,000 to \$20,000 (q7=4, valid skips excluded from base)	10%	10%	8%		
School year costs over \$20,000 (q7=5, valid skips excluded from base)	9%	9%	9%		
Very well prepared for current academic courses (q16=1, valid skips excluded from base)	25%	25%	27%		
Well prepared for current academic courses (q16=2, valid skips excluded from base)	60%	60%	56%		
Not well prepared for current academic courses (q16=3, valid skips excluded from base)	13%	13%	15%		
Very poorly prepared for current academic courses (q16=4, valid skips excluded from base)	2%	2%	3%		

	Total	First 90% of cases	Last 10% of cases	Bias using first 90%	Bias as % of estimate
Wish had selected another college – Agree (q21c=0 to 3, valid skips excluded from base)	68%	69%	65%		
Wish had selected another college – Neutral (q21c=4 to 7, valid skips excluded from base)	20%	19%	21%		
Wish had selected another college – Disagree (q21c=8 to 10, valid skips excluded from base)	12%	12%	14%		
Difficult to graduate in four years – Agree (q21f=0 to 3, valid skips excluded from base)	61%	61%	61%		
Difficult to graduate in four years – Neutral (q21f=4 to 7, valid skips excluded from base)	23%	24%	21%		
Difficult to graduate in four years – Disagree (q21f=8 to 10, valid skips excluded from base)	16%	16%	19%		
Worked for pay since end of senior year (q2b=yes)	79%	79%	83%		
Currently working for pay (q2b_1=yes)	69%	69%	68%		
Parents often checked if homework was done (q27a=3)	31%	31%	37%	-0.6%	-1.9%
Parents sometimes checked if homework was done (q27a=2)	25%	25%	23%		
Parents rarely checked if homework was done (q27a=1)	44%	44%	40%		
Parents often rewarded for good grades (q27d=3)	34%	33%	41%	-0.8%	-2.4%
Parents sometimes rewarded for good grades (q27d=2)	33%	34%	29%	0.5%	1.5%
Parents rarely rewarded for good grades (q27d=1)	33%	34%	30%		
My family often did fun things together (q28f=3)	44%	44%	50%	-0.6%	-1.4%
My family sometimes did fun things together (q28f=2)	38%	39%	35%		
My family rarely did fun things together (q28f=1)	17%	18%	15%		
Had a curfew for school nights (q29=yes)	55%	55%	58%		
I feel good about myself - Strongly agree (q33a=1)	47%	47%	50%		
I feel good about myself - Agree (q33a=2)	51%	52%	48%		
I feel good about myself - Disagree (q33a=3)	1%	1%	1%		
I feel good about myself - Strongly disagree (q33a=4)	0%	0%	0%		
I feel I do not have much to be proud of - Strongly agree (q33i=1)	1%	1%	1%		
I feel I do not have much to be proud of - Agree (q33i=2)	9%	8%	12%		
I feel I do not have much to be proud of - Disagree (q33i=3)	61%	62%	56%	0.7%	1.1%
I feel I do not have much to be proud of - Strongly disagree (q33i=4)	29%	29%	32%		

## WEIGHTING THE DATA

The analyses outlined in this memorandum are based on unweighted data. The selected variables were also weighted to the TEA population and reviewed. Weighting the data tends to dampen divergences between the earlier and last quartiles or between the first 90% and last 10% of cases collected. For completeness, a table showing weighted data for the last 10% of cases compared to the total and the first 90% of cases is presented in the appendix. Similar to Table 1, bias calculations were made for variables that showed at least a 5 percentage point difference in the estimates using the first 90% and last 10% of cases. As was found in the unweighted data, the nominal biases are all less than one percentage point and the relative biases are all under 3 percentage points. Thus, weighting the data serves only to reinforce the findings from the unweighted analyses.

## COST CONSIDERATIONS

Later cases are much more expensive to complete than the earlier cases. In this study the last quartile of respondents to Wave 2 were three times more expensive, and the last 10% were four times more expensive than the first quartile.

TABLE 3: COST TO COMPLETE MEASURED BY COMPLETES PER HOUR (CPH)

<b>Segment</b>	<b>CPH</b>
First 25% of Cases	1.20
Middle 50% of Cases	0.98
Last 25% of Cases	0.43
Last 10% of Cases	0.32
<b>Total</b>	<b>0.80</b>

## WAVE 2 NON-RESPONDERS

Baseline survey data collected at Wave 1 in the spring of 2002 was can be used to provide insight into Wave 2 nonresponders. Reviewing distributions for over 100 variables, the majority show differences between Wave 2 responders and nonresponders of under 2% and no differences were found over 3%. These results contrast with findings of others (e.g., Kennickell (1999), Bates & Creighton (2000), and Chui et al (2001) among others) and contraindicate adopting a 'continuum of resistance' model of nonresponse (e.g., I-Fen & Schaeffer, 1995). This may be due to the special population of adolescents entering adulthood. Another possibility is that it may be an inevitable outcome of a panel survey. Alternatively, it may become evident at Wave 3 that as adults, these responders may behave in line with 'continuum' theory. Additional assessment of the baseline data at various waves of the survey is warranted.

## CONCLUSIONS

Overall we have found that the data vary little between cases that were completed early in data collection and those completed later. Demographic variables such as race/ethnicity and gender are stable throughout data collection.

The pattern of responses to variables selected for this analysis is similar enough across the various time periods of data collection for Wave 2 to suggest that late responders do not differ significantly from earlier responders, be they the last 25 % or the last 10% to participate in the survey.

Reviewing Wave 1 data for Wave 2 responders and nonresponders, variations were minimal and none were above 3%.

Thus, the evidence indicates that data collection could have stopped at 90% or even 75% of Wave 2 target without risk of meaningful bias. Therefore, the marginal cost of adding additional cases at subsequent waves of the survey should be weighed against the value they may or may not provide.

# APPENDIX

## WEIGHTING THE DATA

The analyses outlined in this memorandum are based on unweighted data. The selected variables were also weighted to the TEA population and reviewed. Weighting the data tends to dampen divergences between the earlier and last quartiles or between the first 90% and last 10% of cases collected. For completeness, a table showing weighted data for the last 10% of cases compared to the total and the first 90% of cases is presented in the appendix. Similar to Table 1, bias calculations were made for variables that showed at least a 5 percentage point difference in the estimates using the first 90% and last 10% of cases. As was found in the unweighted data, the nominal biases are all less than one percentage point and the relative biases are all under 3 percentage points. Thus, weighting the data serves only to reinforce the findings from the unweighted analyses.

TABLE 1: LAST 10% OF CASES COMPARED TO TOTAL AND FIRST 90% WEIGHTED TO TEA POPULATION

[WEIGHTED TOTAL N=206,960]

Column 1	2	3	4	5	6
	Total	First 90% of cases	Last 10% of cases	Bias using first 90% <sup>8</sup>	Bias as % of estimate <sup>9</sup>
Date of interview	4/3 to 12/27	4/3 to 8/28	8/29 to 12/27		
WEIGHTED to TEA Population	n=206,960	n=197,238	n=9,722		
<b>Variable</b>					
Race/Ethnicity – White	49%	49%	51%		
Race/Ethnicity - African American	10%	10%	13%		
Race/Ethnicity - Mexican American	28%	28%	27%		
Race/Ethnicity - Other Hispanic	5%	5%	6%		
Race/Ethnicity – Asian	4%	4%	2%		
Race/Ethnicity - Native American	0%	0%	0%		
Race/Ethnicity - Something Else	4%	4%	1%		
Gender – Male	48%	48%	49%		
Gender - Female	52%	52%	51%		
Father born in United States (q39=yes)	75%	75%	83%	0.3%	0.4%
Mother born in United States (q40=yes)	75%	75%	79%		
Graduated from high school (q1=yes)	99%	99%	99%		

<sup>8</sup> “Bias Using First 90%” was calculated if the difference between the result for the last 10% of cases and the result for the first 90% of cases is 5 percentage points or greater.

<sup>9</sup> “Bias as % of Estimate” was calculated if the difference between the result for the last 10% of cases and the result for the first 90% of cases is 5 percentage points or greater.

Attended school since graduating from high school (q2e=yes)	74%	74%	75%		
Currently attending school (q2e_0=yes, valid skips excluded from base)	91%	92%	72%	0.9%	1.0%
Attending school full time (q2e_4=full time, valid skips excluded from base)	88%	88%	90%		
Granted admission through Top 10 Percent Law (q5a=1, valid skips excluded from base)	17%	17%	10%	0.4%	2.4%
	<b>Total</b>	<b>First 90% of cases</b>	<b>Last 10% of cases</b>	<b>Bias using first 90%</b>	<b>Bias as % of estimate</b>
Decided to attend Principal Institution because of Top Ten Law (q6i=yes, valid skips excluded from base)	11%	11%	10%		
Pursuing Bachelor's degree (q2e_6=3, valid skips excluded from base)	63%	63%	65%		
Pursuing Associate's degree (q2e_6=2, valid skips excluded from base)	17%	17%	15%		
School year costs under \$5,000 (q7=1, valid skips excluded from base)	29%	29%	26%		
School year costs \$5,000 to \$10,000 (q7=2, valid skips excluded from base)	32%	32%	36%		
School year costs \$10,000 to \$15,000 (q7=3, valid skips excluded from base)	19%	19%	22%		
School year costs \$15,000 to \$20,000 (q7=4, valid skips excluded from base)	11%	11%	9%		
School year costs over \$20,000 (q7=5, valid skips excluded from base)	10%	10%	7%		
Very well prepared for current academic courses (q16=1, valid skips excluded from base)	26%	26%	27%		
Well prepared for current academic courses (q16=2, valid skips excluded from base)	58%	59%	53%	0.3%	0.5%
Not well prepared for current academic courses (q16=3, valid skips excluded from base)	13%	13%	16%		
Very poorly prepared for current academic courses (q16=4, valid skips excluded from base)	3%	3%	4%		
Wish had selected another college – Agree (q21c=0 to 3, valid skips excluded from base)	70%	71%	67%		
Wish had selected another college – Neutral (q21c=4 to 5, valid skips excluded from base)	18%	18%	16%		
Wish had selected another college – Disagree (q21c=8 to 10, valid skips excluded from base)	11%	11%	17%	-0.3%	-2.4%



Difficult to graduate in four years – Agree (q21f=0 to 3, valid skips excluded from base)	61%	61%	64%		
Difficult to graduate in four years – Neutral (q21f=4 to 7, valid skips excluded from base)	23%	23%	19%		
Difficult to graduate in four years – Disagree (q21f=8 to 10, valid skips excluded from base)	16%	16%	17%		
Worked for pay since end of senior year (q2b=yes)	81%	81%	83%		
Currently working for pay (q2b_1=yes)	69%	68%	72%		
	<b>Total</b>	<b>First 90% of cases</b>	<b>Last 10% of cases</b>	<b>Bias using first 90%</b>	<b>Bias as % of estimate</b>
Parents often checked if homework was done (q27a=3)	31%	31%	42%	-0.5%	-1.6%
Parents sometimes checked if homework was done (q27a=2)	25%	25%	23%		
Parents rarely checked if homework was done (q27a=1)	44%	45%	36%	0.5%	1.1%
Parents often rewarded for good grades (q27d=3)	34%	34%	40%	-0.3%	-0.8%
Parents sometimes rewarded for good grades (q27d=2)	33%	33%	30%		
Parents rarely rewarded for good grades (q27d=1)	33%	33%	30%		
My family often did fun things together (q28f=3)	45%	45%	51%	-0.3%	-0.7%
My family sometimes did fun things together (q28f=2)	39%	39%	36%		
My family rarely did fun things together (q28f=1)	16%	16%	13%		
Had a curfew for school nights (q29=yes)	56%	56%	58%		
I feel good about myself - Strongly agree (q33a=1)	46%	46%	53%	-0.3%	-0.6%
I feel good about myself - Agree (q33a=2)	52%	53%	46%	0.3%	0.6%
I feel good about myself - Disagree (q33a=3)	1%	1%	1%		
I feel good about myself - Strongly disagree (q33a=4)	0%	0%	0%		
I feel I do not have much to be proud of - Strongly agree (q33i=1)	1%	1%	1%		
I feel I do not have much to be proud of - Agree (q33i=2)	9%	9%	11%		
I feel I do not have much to be proud of - Disagree (q33i=3)	62%	63%	50%	0.6%	1.0%
I feel I do not have much to be proud of - Strongly disagree (q33i=4)	28%	28%	38%	-0.5%	-1.8%