Subject versus Population Characteristics in Economics Experiments*

John C. Ham
Department of Economics
University of Southern California, the Federal Reserve Bank of San Francisco and IZA

John H. Kagel
Department of Economics
Ohio State University

3/13/2006

Abstract

This note compares the demographic and ability characteristics from a large sample of experimental subjects to the University population from which they were drawn. Economics and business majors are overrepresented compared to the University population as subjects were recruited from economics classes. However, seventy percent of the sample population consists of science, engineering and other majors. Our sample has a much higher ability level as measured by SAT/ACT scores than the University population as a whole, or than the University population after conditioning on academic major or gender.

Key words: subject characteristics, economics experiments, population characteristics, SAT/ACT scores
JEL classification: C8, C9.

* We thank Susan Rose for able research assistance and helpful comments, Robert Slonim for helpful comments, and Robert Vandyne, Jr. from Student Enrollment and Research Services at the Ohio State University, for his help in obtaining detailed demographic and ability data on students. This research has been supported by NSF Grant 0136928. Any opinions, findings, and conclusions or recommendations in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation, the Federal Reserve Bank of San Francisco or the Federal Reserve System. We are responsible for any errors.

We recently completed a relatively large scale laboratory experiment investigation of bidding in common value auctions (Casari, Ham and Kagel, 2005; hereafter CHK). A key part of this study was to collect ability and demographic data for the subjects to determine their impact, if any, on bidding. The purpose of this note is to compare our sample characteristics with the University population from which they were drawn, as we unaware of any prior studies indicating how representative student subject populations are in economics experiments compared to the University population from which they are drawn. Our key findings are that men and women are recruited in roughly representative proportions to the University population and that experimental subjects have considerably higher cognitive/academic abilities than the University population as measured by SAT/ACT scores.[1] Although it remains to be seen if these results generalize to other academic settings where economics experiments are typically performed, the present results provide a first glimpse into this issue.

Subjects were recruited by e-mail based on enrollment lists from all students registered in undergraduate economics classes in the Fall and Winter quarters of 2002 at the Ohio State University.[2] Enrollment was not limited to such students, and we know from past experience that at least some students (and occasionally non-students) who have not been contacted directly via e-mail participate in the experiments based on word of mouth. As part of the recruitment protocol, prospective subjects were informed that we would ask them to sign a consent form permitting us to access their academic records from the University records office, with the proviso that these records would be used without any ability to identify individual students. The consent form called for providing

[1] As is typical of most economics experiments subjects in this experiment were volunteers who received financial compensation conditional on their performance in the experiment. They are *not* required to participate in the experiments as part of course participation or extra credit.

[2] These enrollment lists contain approximately 5000 names each quarter. They contain a rather broad cross-section of undergraduate students as there are limited numbers of economics majors (Ohio State has a large undergraduate business school), with most of the teaching involving service courses outside the department.

information regarding gender, SAT and ACT scores, major, and class standing (freshman, sophomore, etc.). The information was provided in machine-readable form by the University and was matched to the experimental data through social security numbers. This records requirement did not result in any noticeable impact on our ability to recruit from the same subject population compared to other experiments that did not request such information.

A total of 258 subjects participated in the experiment of which 10 were graduate students, with the remaining subjects all undergraduates at Ohio State University. In what follows we drop the graduate students and compare our sample of undergraduates with the undergraduate University population from which they were drawn.[3]

The demographic and ability measures employed in the data analysis and reported on here are as follows:

Gender: Male/Female.

College major: Three categories were established - business and economics, science and engineering, and a residual category. Language, other social science and military academy majors accounted for about three quarters of the subjects in the latter category. It also included subjects in arts, dance and education as well as undeclared majors.[4]

SAT/ACT scores: These scores result from standardized tests that most high school graduates take when seeking admission to a US college. A set of binary variables was constructed for three ability levels - high, median, and low - based on both SAT and ACT scores. The cut-off points chosen were below the median, between the median and the

[3] The sample employed here differs slightly form the sample reported in CHK as CHK (i) include graduate students and (ii) exclude subjects who were bystanders in some of the initial experimental sessions and never got to bid. Note we do not have information on the handful of "extra" subjects who were sent home and did not participate in a session. Extras were determined randomly at the start of a session after first determining which subjects could return for a later session (with the guarantee that they would get participate upon returning).

[4] See http://www.econ.ohio-state.edu/kagel/majorsclassification for how academic majors were classified. Students declare their major in the first quarter of attending Ohio State University. However, they can essentially declare that they are undecided through the major "exploration in…" which would place then in the "other" category.

95-percentile, and the 95-percentile or higher, where the scores are calibrated relative to the national distribution of scores for the years 2002-2003.[5] Binary variables were generated because (i) ACT and SAT scores are not additive measures of ability but rather rank order measures so that binary scores are more appropriate to use than raw scores and (ii) a number of students were missing SAT or ACT scores (39.4% SAT and 25.5% ACT), which would have increased the fraction of observations with missing values. Using these binary variables permits us to get around this fact by counting someone in the top 5% if she was in the top 5% of the national distribution for *either* SAT or ACT, below the median if they were below the national median in *either* test, and in the middle group otherwise.[6] Using both scores this way reduces the proportion of subjects with a missing value to 13.7% of the sample.[7] Bidders were coded according to their verbal, mathematical, and combined skills. This is possible because the ACT test has separate sub-scores for Reading, English, Mathematics, and Science and Reasoning abilities, while the SAT test has separate sub-scores for Verbal and Mathematics abilities. The verbal ability measure is derived from ACT Reading (ACT English deals with grammar questions) and SAT Verbal, while the combined ability is based on ACT and SAT Composite test scores.[8] Although the categories used for SAT/ACT scores are somewhat arbitrary, they do provide reasonable measures of high, medium and low ability that we believe are interesting.

[5] The cut-off points employed are the following. Type of score (at or below median, at or above 95-centile): SAT Verbal (500,700), Mathematics (510, 700), Combination (505, 700); ACT English (20, 30), Reading (21, 32), Mathematics (19, 30), Science and Reasoning (20, 29), Combination (20, 30). For ACT the reference points are the scores reported during years 2002-2003. For updates on the distribution see http://www.act.org/aap/scores/. For SAT I test the reference are the scores of 2002 distribution. See www.collegeboard.com for more information.

[6] Many students take the SAT or ACT test more than once. In these cases we used the latest score reported.

[7] Most of the subjects with missing values were students who transferred to OSU after the first year from satellite campuses as these scores are not required for these students.

[8] Correlations between SAT and ACT sub-scores are rather high: 0.85 for Mathematics,; 0.76 for Verbal, and 0.86 for Composite.

Grade point average: We also explored a number of empirical specifications using grade point average (GPA) in place of, or in conjunction with, SAT/ACT scores to measure ability (here too using a binary specification). GPA proved to be a far inferior ability measure compared to SAT/ACT scores in CHK, rarely achieving statistical significance in any of the bid function specifications or in the analysis of bankruptcy rates in common value auctions. We suspect there are two primary reasons for this. First, we have a number of freshmen and sophomores in our sample for which GPA would be a very incomplete measure of academic performance. Second, there is likely to be a good deal of heterogeneity in grade scales within our specified college majors, no less between the different majors. The two together make GPA a much fuzzier measure of ability than SAT/ACT scores. We nevertheless include information about grade point averages for completeness sake.

The first two columns of data in Table 1 provide an overall comparison of the sample data versus the undergraduate University population as a whole. There are more men in the sample than in the University population, with this difference statistically significant at conventional levels ($p < .05$). For academic majors, our sample included a much larger percentage of economics and business majors than the University population (30.2% versus 12.3%), which was to be expected given that we recruited based on e-mail lists for students enrolled in economics classes. The proportion of engineering and science majors in our sample (23.4%) almost exactly matches the University population (25.8%). We had less of the residual category, other majors, than the University population.

The sample population includes students with significantly higher ability than the University population as measured by SAT/ACT scores. Focusing on composite scores, the sample had 20.2% in the top 5% versus the University population of 4.9%, with

comparable differences for math and verbal scores. Further, we had a substantially smaller percentage of students in the sample with below median SAT/ACT scores (8.9% versus 20.9% for composite scores). Students with no SAT/ACT scores are primarily transfers from regional campuses as these students are not required to take these tests when transferring to the main campus. Here too we have a smaller percentage than the University population. It seems likely that these transfers would have lower SAT/ACT scores if we had their test records as a number of these regional campus transfers were ineligible to enter to main campus at the time that they applied. Assuming that these transfer students SAT/ACT scores would have put them in the below median category (which no doubt is an extreme assumption), 22.2% of the sample were in the below median category versus 42.0% for the University population ($p < 0.01$).

Grade point averages – computed only for Junior and Seniors – would seem to reflect the higher SAT/ACT scores in the sample population as these students with B averages or above are 32.3% in the sample population versus 24.2% for the University population. The percentage of students with no GPA computed or Freshmen or Sophomores does not differ much between the sample and the University population.[9]

The sample and population statistics for men and women are in columns 3-4 and 5-6 respectively. Consistent with the overrepresentation of economics and business majors in the sample, we have both more men *and* women economics and business majors in the sample than men and women in the University population. (See Table 2 for gender composition in our sample versus the University population, conditioning on major.) The percentage of male engineering and science majors in the sample closes matches the percentage of male science and engineering majors, while the percentage of female science and engineering majors in the sample is underrepresented ($p < .05$). The

[9] In our analysis we did not examine GPAs for Freshmen and Sophomores on the grounds that they had spent too little time at the University for these scores to be meaningful.

percentage of female "other" majors closely matches the University population, while the percentage of male "other" majors is underrepresented in the sample.

The sample population has higher ability levels for both men and women than the corresponding University populations as measured by composite SAT/ACT scores in the top 5% . For men, 25.7% of the sample are in the top 5% versus only 5.9% in the university population ($p< .01$), while for women, 12% of the sample are in the top 5% versus 3.9% in the population ($p < .01$). Thus, the sorting in terms of high ability types into the experimental sample is even more extreme for men than women.[10] Most of this within sample difference comes from the SAT/ACT math scores with 35.8% of the sample men in the top 5% for math (compared to 10.6% in the population) versus 9.0% of the sample women (compared to 4.8% of the population).[11] Note that in the statistical analysis of bidding in the auction experiment itself, it is higher composite SAT/ACT scores that are consistently associated with better performance (lower bankruptcy rates and a larger bid factor) as opposed to math or verbal scores alone.[12] Further, the biggest and most consistent impact of ability comes as a result of those with below median (or missing) SAT/ACT scores being more susceptible to the winner's curse, as opposed to those with very high scores doing exceptionally well, and on this measure there are minimal differences between men and women within the experimental sample. Specifically 19.5% of the sample men are in this combined group (compared with 39.6% of the population) versus 26.0% of the sample women (as compared to 44.5% of the population). Finally, the grade point average data for the sample population differ little

[10] Taking the difference between the proportion of men in the top 5% on composite SAT/ACT scores in our sample versus the male University population and comparing it to the difference between the women in our sample and the female University population in the top 5% yields a difference of 11.7% ($p < .05$).
[11] Taking the difference between the proportion of men in the top 5% on math SAT/ACT scores in our sample versus the male University population and comparing it to the difference between the women in our sample and the female University population in the top 5% yields a difference of 20.9% ($p < .01$).

[12] Bid factor is defined as the private information signal regarding the true value provided to each bidder less their bid (see CHK for details).

between men and women, and at least in terms of the fraction of students receiving at B-average or less, both the male and female experimental participants have better grades than the university population.

Table 2 compares the experimental sample with the University population conditional on a students' academic major. Columns 1 and 2 show the results for economics and business majors. The proportion of male and female students in the sample population does not differ significantly from their proportion in the University population of economics and business majors. Composite SAT/ACT scores show the same pattern as reported earlier, with students with scores in the top 5% significantly more likely to volunteer for the experiment than the University population for economics and business majors (20.0% versus 4.2%; $p < .01$), and substantially fewer students with below median or missing SAT/ACT scores volunteering than the University population (13.3% versus 36.3% pooling these two categories, $p < .01$).

Columns 3 and 4 of Table 2 report the results for science and engineering majors. There are significantly more male than female science and engineering majors in the sample population compared to the University population of science and engineering majors ($p < .05$). As with the sample population as a whole, among science and engineering majors, the experiment attracts a much brighter cohort than the University population. This holds for both men and women science and engineering majors.[13] Grade point averages for science and engineering majors in the experimental sample are also higher than for science and engineering majors as a whole as well.

The last two columns of Table 2 contain our catchall category of other majors. Here too the sample composition of men and women closely matches the University population. Composite SAT/ACT scores are once again better than the University

[13] Top 5% composite SAT/ACT scores: Men in the sample - 37.5% vs University population of men in science and engineering (10.3%, $p < .01$). Women in the sample - 40.0% vs University population of women in science and engineering (8.8%, $p < .01$).

population of such students as the sample had 11.3% of other majors in the top 5% versus 3.0% for the University population ($p < .01$). However, the difference between the sample and University population is somewhat smaller for the top 5% than for the economics and business majors or science and engineering majors, as is the difference in students below the median or with a missing score. Finally, grade point averages are somewhat better for the experimental sample than for the University population of "other" majors as those with GPA of B minus or below account for 13.9% of the sample versus 29.1% for the population of other majors ($p< .01$).

The fact that better students participate in the experiments raises the issue of whether the better students are on our mailing list, or whether they are more likely to respond to our email. For those outside economics and business, we cannot address this question. However for the economics and business students, it is clear that the better students are more likely to respond. The economics department provides introductory economics classes for virtually all students classified as economics and business majors, so that virtually all of these students are on our mailing list. Thus, the higher ability of economics and business students in our sample as compared to the population of economics and business students must result from the higher ability students being more likely to respond to our email. Note that better ability students did better in our auction experiment, and thus it was rational for the better students to be more responsive to our email. Further, this is consistent with the evidence in CHK that the better ability students were more likely to come back to week 2 of the experiment after participating in week 1.

Summary and Conclusions: This note compares the demographic and ability characteristics from a large sample of experimental subjects to the University population from which they were drawn. Economics and business majors are overrepresented compared to the University population as subjects were recruited from economics classes.

Although the sample and population gender proportions are statistically significantly different from each other, the overall proportions are not terribly different (59.7% men in the sample versus 51.8% in the University population). Our sample has a much higher ability level as measured by SAT/ACT scores than the University population as a whole, or than the University population conditional on academic major or gender.

These results with respect to SAT/ACT scores are perhaps the most surprising of the results reported. While the present study does not directly compare our experimental subjects with the more general population, 19.9% of experimental sample had composite SAT/ACT scores in the top 5% which is well above the percentage taking the tests (5%) in the University population from which they were drawn (4.9%). Further, the experimental sample contains a far smaller percentage of students with below median SAT/ACT scores than the sample of students taking these tests as a whole or than the University population, even assuming that all those with no SAT/ACT scores (primarily transfer students) would have had below median SAT/ACT scores. Relating these results to the results for the experiment conducted, the biggest and most consistent impact of ability comes as a result of those with below median scores being more susceptible to the winner's curse, as opposed to those with very high scores doing exceptionally well.

There is a broader question lurking in the background here that our results serve to address. Models of economic behavior typically postulate "rational" economic behavior, which is presumably positively related to IQ, or ability, for which SAT/ACT scores serve as a reasonable proxy. When we observe behavior in experiments that does not match our models – for example, the winner's curse in common value auctions or relatively slow emergence of equilibria in signaling games, or equilibria that violate even the most basic equilibrium refinement criteria (see, for example, Brandts and Holt, 1992 or Cooper, Garvin and Kagel, 1997) – the question becomes how representative are the

results likely to be as compared to those for the general population? To the extent that rational behavior is positively related to IQ or intellectual ability as measured by SAT/ACT scores our experimental sample is drawing from the upper end of the distribution on this account, suggesting that in general economics experiments, with their financially motivated volunteer subjects, attract the upper end of the University population from which they are drawn. Although this does not imply that we are dealing with people who are as "smart" or "expert" as one can imagine in a large number of field settings (e.g., oil company personnel bidding for oil exploration rights or a signaling game involving firm behavior), for a variety of other applications we are no doubt dealing with a much "smarter" cohort in the lab than the general population (e.g., applications of the winner's curse to jury decisions or attempts by corporations to use advertising to signal product quality to the general population). The results here make it much harder to reject results from experiments that violate one's pet model, at least to the extent that the model is applied to general populations, on the grounds that the student subjects are just not "smart enough."

References

Brandts, Jordi, and Holt, Charles A. "An Experimental Test of Equilibrium Dominance in Signaling Games," American Economic Review, December 1992, 82(5), 1350-65.

Casari, Marco, Ham, John C. and Kagel, John H. "Selection Effects, Demographic Effects and Ability Effects in Common Value Auction Experiments," Working Paper, Ohio State University (2005).

Cooper, David, J., Garvin, Susan, and Kagel, John H. "Adaptive Learning and Equilibrium Refinements in an Entry Limit Pricing Game," Economic Journal, May 1997, 107(442), 553-75.

Table 1
Sample versus Population Differences: Overall and by Gender
(Percentages and Chi-Square Values)

	All		Male		Female	
Gender	Sample	Population	Sample	Population	Sample	Population
Male	59.7	51.8				
Female	40.3	48.2				
	$\chi^2(1) = 6.15^{**}$					
Major						
Econ & Bus	30.2	12.3	35.1	14.3	23.0	10.2
Eng & Science	23.4	25.8	32.4	33.7	10.0	17.3
Other	46.4	61.9	32.4	52.0	67.0	72.6
	$\chi^2(2) = 74.7^{***}$		$\chi^2(2) = 55.8^{***}$		$\chi^2(2) = 19.5^{***}$	
SAT/ACT Composite						
Top 5%	20.2	4.9	25.7	5.9	12.0	3.9
Above median but not top 5%	57.7	53.2	54.7	54.5	62.0	51.6
Below median	8.9	20.9	5.4	17.8	14.0	24.2
No score	13.3	21.1	14.2	21.8	12.0	20.3
	$\chi^2(3) = 143.5^{***}$		$\chi^2(3) = 116.0^{***}$		$\chi^2(3) = 26.9^{***}$	
SAT/ACT Math						
Top 5%	25.0	7.8	35.8	10.6	9.0	4.8
Above median but not top 5%	52.0	49.8	46.0	52.2	61.0	47.2
Below median	9.7	21.3	4.1	15.4	18.0	27.8
No score	13.3	21.1	14.2	21.8	12.0	20.3
	$\chi^2(3) = 117.0^{***}$		$\chi^2(3) = 105.6^{***}$		$\chi^2(3) = 14.6^{***}$	
SAT/ACT Verbal						
Top 5%	17.7	7.0	18.2	6.9	17.0	7.0
Above median but not top 5%	51.2	43.9	53.4	44.0	48.0	43.7
Below median	17.7	28.1	14.2	27.3	23.0	28.9
No score	13.3	21.1	14.2	21.8	12.0	20.3
	$\chi^2(3) = 60.7^{***}$		$\chi^2(3) = 43.6^{***}$		$\chi^2(3) = 19.1^{***}$	
Grade Point Average						
A	6.5	4.7	6.8	4.0	6.0	5.5
B+ / B	25.8	19.5	27.7	17.5	23.0	21.7
B- or below	12.5	27.6	12.8	30.5	12.0	24.4
Freshman, Sophomore or no GPA	55.2	48.2	52.7	48.0	59.0	48.4
	$\chi^2(3) = 29.6^{***}$		$\chi^2(3) = 27.5^{***}$		$\chi^2(3) = 8.8^{**}$	

* Significantly different from zero at 10% level
** Significantly different from zero at 5% level
*** Significantly different from zero at 1% level

Table 2
Sample versus Population Differences by Academic Major
(Percentages and Chi-Square Values)

	Economics and Business		Science and Engineering		Other	
Gender	Sample	Population	Sample	Population	Sample	Population
Male	69.3	60.1	82.8	67.8	41.7	43.5
Female	30.7	39.9	17.2	32.2	58.3	56.5
	$\chi^2(1) = 2.65$		$\chi^2(1) = 6.0^{**}$		$\chi^2(1) = 0.1$	
SAT/ACT Composite						
Top 5%	20.0	4.2	37.9	9.8	11.3	3.0
Above median but not top 5%	66.7	59.5	58.6	60.5	51.3	48.8
Below median	9.3	16.2	1.7	12.3	12.2	25.4
No score	4.0	20.1	1.7	17.4	25.2	22.8
	$\chi^2(3) = 56.4^{***}$		$\chi^2(3) = 60.0^{***}$		$\chi^2(3) = 35.3^{***}$	
SAT/ACT Math						
Top 5%	33.3	7.7	48.3	17.5	7.8	3.8
Above median but not top 5%	56.0	58.1	46.6	55.2	52.2	45.8
Below median	6.7	14.2	3.5	9.9	14.8	27.5
No score	4.0	20.1	1.7	17.4	25.2	22.8
	$\chi^2(3) = 77.3^{***}$		$\chi^2(3) = 42.9^{***}$		$\chi^2(3) = 13.0^{***}$	
SAT/ACT Verbal						
Top 5%	21.3	5.9	20.7	10.6	13.9	5.7
Above median but not top 5%	57.3	47.1	62.1	49.1	41.7	41.0
Below median	17.3	26.9	15.5	23.0	19.1	30.4
No score	4.0	20.1	1.7	17.4	25.2	22.8
	$\chi^2(3) = 43.9^{***}$		$\chi^2(3) = 17.2^{***}$		$\chi^2(3) = 18.8^{***}$	
Grade Point Average						
A	10.7	6.7	6.9	5.7	3.5	3.9
B+ / B	28.0	32.8	34.5	19.6	20.0	16.9
B- or below	6.7	24.0	17.2	25.8	13.9	29.1
Freshman, Sophomore or no GPA	54.7	36.5	41.4	48.9	62.6	50.2
	$\chi^2(3) = 18.5^{***}$		$\chi^2(3) = 9.0^{**}$		$\chi^2(3) = 13.4^{***}$	

* Significantly different from zero at 10% level
** Significantly different from 0 at the 5% level
*** Significantly different from zero at 1% level