TEST OF UNDERSTANDING IN COLLEGE ECONOMICS

Fourth Edition

Examiner's Manual

William B. Walstad
Michael Watts
Ken Rebeck

National Council on Economic Education

Test of Understanding of College Economics (Fourth Edition) Examiner's Manual

William B. Walstad

Michael Watts

Ken Rebeck

(July 24, 2006)

ABOUT THE AUTHORS

William B. Walstad is Professor of Economics at the University of Nebraska-Lincoln, Director of the National Center for Research in Economic Education, and a past Chair of the Committee on Economic Education of the American Economic Association.

Michael W. Watts is Professor of Economics at Purdue University, Director of the Purdue Center for Economic Education, and Chair of the Committee on Economic Education of the American Economic Association.

Ken Rebeck is Associate Professor of Economics at St. Cloud State University (Minnesota) and a Research Associate with the National Center for Research in Economic Education.

ISBN 1-56183-609-5 5 4 3 2 1

TABLE OF CONTENTS

TABLE OF CONTENTS
(continued)

Page

FOREWORD

The National Council on Economic Education (NCEE) is deeply committed to providing the highest quality products for instructors to use in their classrooms to help students learn economics. Publishing up-to-date assessment instruments is part of the important contribution the NCEE makes to advancing economic education at the pre-college and college levels. This Examiner's Manual for the fourth edition of the *Test of Understanding of College Economics* provides the test administrator with information on giving the test to students and provides the instructor with information to compare his/her students' performance with that of similar students attending colleges and universities across the nation.

NCEE is truly indebted to many individuals who shared their multitude of talent and precious time to review and revise the questions in this fourth edition of the *Test of Understanding of College Economics*. Special thanks go to William Walstad, Michael Watts, and Ken Rebeck for undertaking and managing this work, and for writing this Examiner's Manual. The members of the Test Development Committee helped prepare and review questions at various stages of the project and assisted with the field testing of items. A National Advisory Committee also reviewed the test. The distinguished members of both committees are acknowledged by name and institution in the Examiner's Manual (on pp. 1-2).

In all and through all, NCEE gratefully acknowledges the generous funding and magnanimous support of the Spencer Foundation for making this accomplishment possible.

Robert F. Duvall, Ph.D.
President and Chief Executive Officer
National Council on Economic Education (NCEE)

THE TEST OF UNDERSTANDING OF COLLEGE ECONOMICS (4TH EDITION): EXAMINER'S MANUAL

This edition of the *Test of Understanding in College Economics* (TUCE-4) is the fourth edition of a test that was first developed forty years ago, and has an extensive history of use by teachers and researchers in the economics profession. The previous editions and their use have been described in earlier studies (Fels, 1967; Welsh and Fels, 1969; Saunders, 1981; Saunders, Fels and Welsh, 1981; Saunders, 1991a; and Saunders, 1991b) and in three reviews of research in economic education at the college and university level (Siegfried and Fels, 1979; Becker, 1997; and Siegfried and Walstad, 1998).

As with past editions, the TUCE-4 has two main objectives: (1) to offer a reliable and valid assessment instrument for students in principles of economics courses; and (2) to provide norming data for a large, national sample of students in principles classes, allowing instructors to compare performance in their classes on both pretests and posttests to the performance of the national sample of students and instructors. Separate exams were prepared in microeconomics and macroeconomics. Both exams consist of 30 multiple-choice items, which can be administered within the time constraints of a single class period for most course formats. The same exams were used for the pretest and posttest, as was done with the third edition of the TUCE. The following sections explain the revision process for the TUCE-4 and provide the results from the national norming of the test with students taking semester-long courses in the principles of microeconomics or macroeconomics.

1. TEST DEVELOPMENT AND REVISION

This revision of the TUCE was once again a joint effort of the Committee on Economic Education of the American Economic Association and the National Council on Economic Education (NCEE), which provided the funding for the TUCE-4 revision through a grant from the Spencer Foundation. The committee members responsible for selecting, writing, and editing the questions on the TUCE-4 were Stephen Buckles, Vanderbilt University; William Bosshardt, Florida Atlantic University; Rae Jean Goodman, U.S. Naval Academy; Paul Grimes, Mississippi State University; Claire Melican (then at the NCEE); William Walstad, University of Nebraska-Lincoln; and Michael Watts, Purdue University. Walstad was the general project director and Watts served as the chair of the revision committee. Melican and Elizabeth Webbink were the NCEE administrators for the project. Ken Rebeck, St. Cloud State University, also reviewed questions and analyzed test data as the associate director for the project. The NCEE recruited instructors and classes to participate in the norming sample, and organized and collected the norming data.

The TUCE-4 Revision Committee began working in spring 2004 with the selection of the committee members made by the project director and the NCEE administrator. The committee prepared content specifications; reviewed, revised or replaced existing test questions; and wrote new questions to fill content gaps. This work was conducted over a 15-month period, entailing four drafts of the two exams. The final draft was nationally normed in the fall semester, 2005.

Content and cognitive specifications were completed in July 2004, and the committee met in August 2004 to produce the first draft of the TUCE-4. After further review and revisions by the committee, a second draft was field-tested as a pretest at the beginning of the spring 2005 semester. The microeconomics test was administered to 660 principles students at six universities, and the macroeconomics test was administered to 1,820 students at seven universities. For comparative purposes, each test was also administered to students taking an intermediate theory course

in microeconomics or macroeconomics (with 40 students in micro and 43 in macro).

The results from the pretest field-testing were analyzed to identify and replace a relatively small number of items with problems. Those revisions yielded a third draft of the TUCE-4 for a "post-test" field-testing conducted at the end of the spring 2005 semester. The third draft of the micro exam was administered to 635 principles students at six universities and the third draft of the macro exam was administered to 1,879 principles students at seven universities.

During this time period, comments on the third draft were provided by a national panel of distinguished economists. The members of this review panel were: Ted Bergstrom, University of California-Santa Barbara; Daniel Hamermesh, University of Texas at Austin; Alan Krueger, Princeton University; W. Douglas McMillin, Louisiana State University; Arthur J. Rolnick, Federal Reserve Bank of Minneapolis; Paul Romer, Stanford University; and Michael Salemi, University of North Carolina-Chapel Hill. Most members of the panel reviewed either the micro or macro exam, but a few reviewed both exams.

The comments from this national panel and the data analysis from the spring posttesting were reviewed by Watts, Walstad, and Melican at a July 2005 meeting. Test items with weak item statistics were eliminated and other questions were deleted or revised to address concerns raised by the national panel, often to strengthen a particular distractor or wording in a question stem. The full TUCE-4 revision committee participated in writing replacement questions and revising these questions. This resulted in the fourth and final draft of the TUCE-4, which was used for the fall 2005 national testing.

2. CONTENT SPECIFICATIONS

The test development committee prepared the following content categories for the microeconomics test and set the following recommended percentage ranges (shown in parentheses) for the allocation of test items.

A. The Basic Economic Problem (scarcity, opportunity cost, choice**)** (10–15%)

B. Markets and Price Determination (determinants of supply and demand, utility, elasticity, price ceilings and floors) (20–25%)

C. Theories of the Firm (revenues, costs, marginal analysis, market structures) (25–30%)

D. Factor Markets (wages, rents, interest, profits, income distribution) (10–15%)

E. The (Microeconomic) Role of Government in a Market Economy (public goods, maintaining competition, externalities, taxation, income redistribution, public choice) (15–20%)

F. International Economics (comparative advantage, trade barriers, exchange rates) (10–15%)

These specification categories are basically the same as those found on TUCE-3, although some of the general descriptions are new, with the older lists of concepts moved to the parenthetical listings of topics for greater format consistency across topics. The stability in general content categories is also reflected in the test items. There are, in fact, only seven entirely new questions on this exam, with 23 items taken from the third edition, though often revised. This "default" position of staying with items from earlier editions, unless there were reasons to change based on minor revisions in the content specifications or problems with item statistics, was explicitly endorsed by the test revision committee. That was done partly because the committee viewed the TUCE-3 micro exam as still generally strong and viable, but also because the time and budget constraints for developing and, especially, field-testing new and substantially revised items were very tight.

As always, it was difficult to find items acceptable to large numbers of economists teaching at different colleges and universities—and which

also exhibited good item statistics. In a few cases we tried new items on the field-test version of the exam, but reverted to the old item (sometimes revised) if the item statistics or comments from external reviewers suggested problems with the new question. There are, however, some questions on new topics, and in new formats, on the micro TUCE-4 exam. For example, there is a question on game theory, which is now covered in virtually every principles of economics textbook. And for the first time ever on the TUCE, one question features a simple graphical model.

An international category is included on both tests. The last three questions on the micro test cover international concepts with a micro orientation (comparative advantage, trade barriers, and exchange rates), while the last three questions on the macro test focus on international concepts with a macro orientation (balance of payments, exchange rate systems, open-economy macro). The committee thought that international concepts are now routinely covered in both principles courses, but recognized that there may be greater variance in the coverage or emphasis given to the international concepts by instructors and textbook authors and publishers.

The content specifications and recommended percentage ranges for the allocation of test items on the macroeconomics exam for the TUCE-4 are:

A. **Measuring Aggregate Economic Performance** (GDP and its components, real vs. nominal values, unemployment, inflation) (10–15%)

B. **Aggregate Supply and Aggregate Demand** (potential GDP, economic growth and productivity, determinants and components of AS and AD, income and expenditure approaches to GDP, the multiplier effect) (25-30%)

C. **Money and Financial Markets** (money, money creation, financial institutions) (10–15%)

D. **Monetary and Fiscal Policies** (tools of monetary policy, automatic and discretionary fiscal policies) (25–30%)

E. **Policy Debates** (policy lags and limitations, rules vs. discretion, long run vs. short run, expectations, sources of macroeconomic instability) (10–15%)

F. **International Economics** (balance of payments, exchange rate systems, open-economy macro) (10–15%)

The macro specifications were revised more because of the greater changes that have occurred in the content and teaching of macroeconomic principles courses since the last revision of the TUCE. For example, aggregate supply and demand models are used in most principles courses and textbooks, but not always, and some recent textbooks written by prominent economists have made a major point in not using them. There has also been some de-emphasis in the coverage of "competing schools" (classical, Keynesian, monetarist, new classical, post-Keynesian, etc.), and in calculating various multipliers. Changes in monetary policy rules and regimes, and in empirical and theoretical models of such topics as economic growth, are also affecting the content of most macro principles courses and textbooks. This revision reflects those changes and whatever content consensus there is for a course on macroeconomic principles. As a result, there are 10 new items on the TUCE-4 macro exam, and extensive revisions on most of the 20 other items taken from the third edition of the TUCE.

Tables 1 and 2 classify each of the 30 micro and macro TUCE-4 questions, respectively, in the six broad content categories identified in the content specifications listed above. The main purpose of these content specifications is to ensure that items on the test cover the core content in a "typical" principles course. If that is done successfully, the total raw score on the exam provides a useful measure of students' general understanding of basic economics principles.

TABLE 1. TUCE-4: Microeconomics Test: Content and Cognitive Specifications

	Cognitive Categories			
Content Categories	**Recognition & Understanding**	**Explicit Application**	**Implicit Application**	**Total (Percent)**
A. Basic Problem		8	10	2 (6.6)
B. Markets & Prices	19	1, 2, 3, 9, 11*, 18		6.5 (21.6)
C. Theories of Firm	4, 12	11*, 14, 17, 21	13, 20, 22	8.5 (28.3)
D. Factor Markets		5	23, 24	3 (10)
E. Micro Role of Government	6, 25, 27	7, 15, 26	16	7 (23.3)
F. International (micro)		28, 30	29	3 (10)
Total (Percent)	6 (20.0)	16 (53.3)	8 (26.6)	30 (100)

Note: A description of each content category is given below. Items with asterisk (*) are allocated .5 to each category.

A. The Basic Economic Problem (scarcity, opportunity cost, choice)

B. Markets and Price Determination (determinants of supply and demand, utility, elasticity, price ceilings and floors)

C. Theories of the Firm (revenues, costs, marginal analysis, market structures)

D. Factor Markets (wages, rents, interest, profits, income distribution)

E. The (Microeconomic) Role of Government in a Market Economy (public goods, maintaining competition, externalities, taxation, income redistribution, public choice)

F. International Economics (comparative advantage, trade barriers, exchange rates)

TABLE 2. TUCE-4: Macroeconomics Test: Content and Cognitive Specifications

	Cognitive Categories			
Content Categories	**Recognition & Understanding**	**Explicit Application**	**Implicit Application**	**Total (Percent)**
A. Measuring Aggregate Performance	1	2, 11, 19		4 (13.3)
B. Aggregate Supply & Demand	4, 17*	3, 13, 14, 21, 23*	15, 20*	7.5 (25.0)
C. Money & Financial Markets	5	12, 16, 22		4 (13.3)
D. Monetary & Fiscal Policies	8, 17*	6, 7, 18, 23*	20*, 24, 25, 27	8.5 (28.3)
E. Policy Debates & Applications	9	10	26	3 (10)
F. International (macro)		30	28, 29	3 (10)
Total (Percent)	6 (20.0)	16 (53.3)	8 (26.6)	30 (100)

Note: A description of each content category is given below. Items with asterisk (*) are allocated .5 to each category.

A. Measuring Aggregate Economic Performance (GDP and its components, real vs. nominal values, unemployment, inflation)

B. Aggregate Supply and Aggregate Demand (potential GDP, economic growth and productivity, determinants and components of AS and AD, income and expenditure approaches to GDP, the multiplier effect)

C. Money and Financial Markets (money, money creation, financial institutions)

D. Monetary and Fiscal Policies (tools of monetary policy, automatic and discretionary fiscal policies,)

E. Policy Debates (policy lags and limitations, rules vs. discretion, long run vs. short run, expectations, sources of macroeconomic instability)

F. International Economics (balance of payments, exchange rate systems, open-economy macro)

Content classifications of individual test items are often difficult to do, however, because questions often cover more than one concept or principle. In cases where the correct alternative deals with a concept or principle in one category and incorrect alternatives deal with concepts or principles in other categories, test items were generally classified in the category corresponding to the correct alternative. For one item on the micro test (#11) and three items on the macro test (#17, #20, and #23), the interaction between the alternatives and the situation posed in the stem was sufficiently complex to justify listing the questions in two different content categories.

Individual questions in each content category vary in difficulty, so no attempt should be made to generalize about the economic understanding of students on a particular concept or principle based on answers to a single question or few questions. It is worth restating that the TUCE-4 is an assessment instrument for measuring the general understanding of principles of economics, not a test of understanding each concept or principle included on the test. Individual instructors or researchers who find that the content specification categories or weightings of these tests are not appropriate for their courses should use the detailed item analysis data discussed below to help interpret their results, or perhaps modify the TUCE-4 exams for use with their students. Modifications will, however, affect the validity and reliability of the test, so that issue should be noted when using the test in a modified form.

3. COGNITIVE SPECIFICATIONS

The Taxonomy of Educational Objectives (Bloom, 1956) is a widely-cited cognitive scheme with six categories: knowledge, comprehension, application, analysis, synthesis, and evaluation. The TUCE-4 uses a modified version of that taxonomy with three broad cognitive categories: Recognition and Understanding (RU), Explicit Application (EA), and Implicit Application (IA). Recognition and Understanding is a combination of Bloom's first two categories. Explicit Application and Implicit Application address one or more of the other three categories (synthesis omitted). The three TUCE-4 cognitive categories are the same as the ones used in the third edition and are defined below.[1]

(RU) Recognizes and Understands Basic Terms, Concepts, and Principles

1.1 Selects the best definition of a given economic term, concept, or principle

1.2 Selects the economic term, concept, or principle that best fits a given definition

1.3 Identifies or associates terms that have closely related meanings

1.4 Recalls or recognizes specific economic rules, e.g., an individual firm's profit is maximized at the level of output at which marginal cost equals marginal revenue

(EA) Explicit Application of Basic Terms, Concepts, and Principles

2.1 Applies economic concepts needed to define or solve a particular problem when the concepts are explicitly mentioned

2.2 Distinguishes between correct and incorrect application of economic concepts that are specifically given

2.3 Distinguishes between probable and improbable outcomes of specific economic actions or proposals involving no unstated assumptions

2.4 Judges the adequacy with which conclusions are supported by data or analysis involving no unstated assumptions

[1]The first edition of TUCE used "Simple Application" and "Complex Application" (Fels, 1967, pp. 664-66) instead of the current "Explicit Application" and "Implicit Application."

(IA) Implicit Application of Basic Terms, Concepts, and Principles

3.1 Applies economic concepts needed to define or solve a particular problem when the concepts are not explicitly mentioned

3.2 Distinguishes between correct and incorrect application of economic concepts that are not specifically given

3.3 Distinguishes between probable and improbable outcomes of specific economic actions or proposals involving unstated assumptions

3.4 Judges the adequacy with which conclusions are supported by data or analysis involving unstated assumptions

Tables 1 and 2 show that 80 percent on each test are application items. This proportion is greater than the 67 percent on the third edition. It is, however, consistent with the general purpose of all previous editions of the TUCE, which have sought to emphasize the application of basic concepts and principles over simple recognition of terms and recall of information. The chair of the original TUCE committee noted: "The test will emphasize the ability to apply economic principles to real problems, including issues of public policy" (Fels, 1967, p. 664).

As with the content categories, classifying test items by cognitive type is not precise. Whether the cognitive processes used by students to answer these questions correspond to the level assigned to each question cannot be known with certainty; and any question for which a student has seen the correct answer can become a recall question, regardless of its classification. Despite these caveats, the main purpose in using the cognitive specifications is to ensure that a large number of questions require application, analysis, or evaluation, not simply recognition and recall. The general goal is for the total score on the TUCE-4 to be a useful measure of students' ability to understand and, even more, apply economic terms, concepts, and principles.

One final point on these classifications is worth noting. There is no direct relation between the difficulty of test items and their cognitive level. Item difficulty, as measured by the percentage of correct responses, can vary across all cognitive levels.

4. TEST DATA

The test data in this manual provide results from a large national sample of college and university students who took a principles of microeconomics or principles of macroeconomics course. Test users can inspect these samples and compare them to the scores of their students when they administer the TUCE-4. In addition, data from this sample were used to interpret the results from particular test items and to evaluate the reliability and validity of the TUCE-4. The sections that follow describe the norming sample and technical characteristics of each test, including means, distribution, and item analysis.

Norming Sample

As shown in Table 3, 5,480 students took the micro TUCE-4 test and 5,517 took the macro TUCE-4 test during the 2005 fall term. These students are divided into three different groups: (1) Most students took the TUCE-4 both as a pretest and a posttest (3,255 micro; 2,789 macro). (2) Some students took the TUCE-4 only as a pretest (1,621 micro; 2,022 macro) because instructors ran out of class time to give the posttest or decided not to administer it for other reasons, and because some students dropped the course or were absent at the time of the posttest. (3) Some students only took the posttest (604 micro; 706 macro) because some instructors who did not administer the pretest decided to administer the posttest or were added for the posttest, and because students who were absent for the pretest took the posttest or transferred into the section after the pretest was administered.

TABLE 3. Aggregate Statistics for TUCE-4 Norming Sample

	Micro	Macro
Total Tested		
Matched (pre & post)	3,255	2,789
Pretest only	1,621	2,022
Posttest only	604	706
Total	5,480	5,517
Samples		
Matched (pre & post)		
Students	3,255	2,789
Institutions	43	44
Instructors	71	62
Unmatched (pre & post)		
Pretest total		
Students	4,876	4,811
Institutions	50	50
Instructors	84	81
Posttest total		
Students	3,859	3,495
Institutions	44	46
Instructors	72	64
Mean Scores		
Matched		
Pretest	9.39	9.80
	(3.32)	(3.48)
Posttest	12.77	14.19
	(4.68)	(5.29)
Change (%)	36%	45%
Unmatched		
Pretest total	9.37	9.76
	(3.35)	(3.48)
Posttest total	12.59	14.06
	(4.68)	(5.28)
Change (%)	34%	44%
Reliability		
Coefficient alpha		
Matched-Post	.70	.77
Unmatched-Post	.70	.77
Standard Error of Measurement		
Matched-Post	2.58	2.53
Unmatched-Post	2.58	2.53

Note: Standard deviations are in parentheses.

The main groups that were used to norm the TUCE-4 were the "matched" samples who took the micro or macro TUCE-4 as a pretest at the beginning and as a posttest at the end of the fall term (3,255 micro; 2,789 macro). These matched samples constitute 50 percent or more of all students who took the micro or macro TUCE-4. The matched group for each test consists of the same students who took the pretest and posttest, so differences in pretest and posttest scores implicitly control for the characteristics of students. These matched samples, therefore, are used for most of the analysis presented here, with most of the data reported in the main tables of this manual.

It is also possible to make group comparisons using the larger *unmatched* groups of all students for each test who either took the pretest or the posttest. This results in sample sizes of 4,876 for the micro pretest (1,621 who took only the pretest plus the 3,255 matched pre-post sample); 3,859 for the micro posttest (604 posttest-only students plus the 3,255 matched pre-post students); 4,811 for the macro pretest (2,022 who took only the pretest plus the 2,789 matched pre-post group); and 3,859 for the macro posttest (704 posttest-only students plus the 3,255) matched pre-post students).

The problem with using these total pretest or total posttest groups for score comparisons is that differences in the composition of the two groups may account for some of the score differences. For this reason, the unmatched samples receive only limited review and analysis, with aggregate results reported in Table 3, but some additional analysis is reported in Appendix 3 (Tables A1–A6). As shown in Table 3, however, there were very small differences in the descriptive statistics for the exams across the matched and total pretest and posttest groups.

Table 3 reports the number of instructors and institutions participating in the national norming of the TUCE-4. For the *matched* samples, micro TUCE-4 data were collected by 71 instructors at 43 institutions and macro TUCE-4 data were collected by 62 instructors at 44 institutions. For the micro test, the number of the same students who

took the pretest and posttest at each institution ranged from 17 to 265, with an average of 79 students. For the macro test, the number of the same students who took the pretest and posttest at each institution ranged from 9 to 339 students, with an average of 63 students.

A coded list of the institutions showing the matched pre-post sample sizes at each institution and institution type is found in Appendix 1. The institution type in the matched samples was defined using Carnegie Foundation classifications (www.carnegiefoundation.org) for the academic degree orientation of the institution. The micro norming sample of 43 institutions included 7 associate's colleges offering degrees, 4 colleges offering only baccalaureate degrees, 25 universities offering up to a master's degree, and 7 doctoral-granting or research universities. The macro norming sample of 44 institutions included 4 associate's colleges offering two-year degrees, 7 colleges offering only a baccalaureate degree, 27 universities offering up to a master's degree, and 6 doctoral-granting or research universities.

The percentage of students in the matched pre-post samples by institution type, as shown in Appendix 1, can be compared with Carnegie Foundation data for 2005 on undergraduate enrollment by institution type. The Carnegie Foundation reports the following distribution: associate's colleges (39%); baccalaureate colleges (8%); master's universities (23%); and doctoral universities (28%). For the micro TUCE-4 sample, the percentages of tested students by institution type were: associate's colleges (7%); baccalaureate college (7%); master's universities (62%); and doctoral (23%). For the macro sample, the percentages are: associate's colleges (7%); baccalaureate colleges (13%); master's universities (53%); and doctoral universities (27%). The TUCE-4 samples show that for baccalaureate colleges (7–13%) and doctoral universities (23–27%) the percentage tested with the TUCE-4 are roughly comparable with the Carnegie percentages for overall undergraduate enrollment. The TUCE-4 samples, however, are over-weighted with students at master's universities and under-weighted with students at associate's colleges, relative to the Carnegie distributions. Such differences may be appropriate because, historically, the TUCE-4 is more likely to be used by economics instructors at master's universities than at associate's colleges. The enrollment weighting is also consistent with practices in norming previous TUCE editions.

Although substantial work was done to obtain a *representative* sample for norming the TUCE-4 across a large national sample of colleges and universities, it should be emphasized that neither the micro sample nor the macro sample is a *random* sample. Some instructors at institutions who were initially asked to participate in the national norming chose not to do so, and some instructors at institutions who agreed to participate failed to provide a complete set of data for a variety of reasons. Nevertheless, the matched student samples from the 43–44 colleges and universities administering the TUCE-4 are broadly representative of the wide range of students taking principles of economics in different types of U.S. institutions of higher education.

Overall Test Results

The mean scores for the micro and macro TUCE-4 reported in Table 3 show that both tests are challenging for the matched pre-post sample of students. The pretest mean was 9.39 on the micro exam and 9.80 on the macro exam. On the pretest, students can correctly answer about 31 percent of the micro test items and 33 percent of the macro test items. These pretest percentages for the TUCE-4 tests are fairly close to a "pure guessing" level of 25 percent for a four-option multiple-choice test. These results mean that there is a substantial range or "headroom" for increasing test scores from pretest to posttest for principles of economics instructors who wish to use the TUCE-4.

By the posttest, mean scores do improve (to 12.77 on micro and 14.19 on macro). The posttest micro score shows that students can correctly answer about 43 percent of the test items whereas

the posttest macro score shows that students can correctly answer about 47 percent of the test items. Compared to classroom tests used by most instructors for grading purposes, the posttest mean scores on the TUCE-4 are low (43–47 percent correct). It is important to emphasize that the TUCE-4 is designed to be a *norm-referenced* test that can be used to discriminate among students across a broad range of intellectual ability and knowledge. To provide appropriate levels of item discrimination and test reliability for research purposes, overall mean scores of around 50 percent are desirable.

The pretest and posttest means are indicative of the results that would be obtained for the typical economics instructors who gave the TUCE-4 as a pretest and posttest in their principles of economics courses. The sample data, however, should not be considered as an absolute standard of achievement in economics but a relative measure. The score can aid economics instructors in comparing their students with others. The comparisons will be meaningful only to the extent that composition of the student body at an institution is similar to the norming sample tested.

The absolute differences in pretest and posttest scores can also be used to calculate the percentage gain from the pretest score. The micro results show a 36 percent increase over the pretest score. The macro results show a 45 percent increase over the pretest scores. The significant improvement in mean scores from pretest to posttest indicates that the TUCE-4 does provide an overall measure of learning in principles of economics.

Also shown in Table 3 are the mean pretest and posttest scores for the unmatched, but larger, samples of students who took the TUCE-4 as a pretest or as a posttest. Those mean scores are almost the exactly same as that for the matched sample. These results were not completely surprising because the matched pre and post sample accounts for 50 percent or more of the pretest total or posttest total samples. For the sake of parsimony, only the matched results will be reported in the remaining tables in the main section of the manual. The manual also focuses on the matched sample because, as previously noted, it provides the most control over student characteristics from pretest to posttest. Some test users, however, might want to know how all students who took a pretest scored even if they did not take a posttest, or how all students scored on a posttest, even if they did not take a pretest. To address those needs, detailed data analysis for the unmatched sample, on the same measures shown for the matched sample, are reported in Appendix 3 (Tables A1–A6).

Percentile Tables

Table 4 presents the raw test scores, their distribution, and the corresponding percentile ranks from the 3,255 college and university students who took the microeconomics version of the TUCE-4 as a pretest and posttest. Table 5 presents the same data for the 2,789 college and university students who took the macroeconomics version of the TUCE-4. The percentile ranks were obtained by calculating the total percentage of students who scored at or below a certain raw score. The pretest percentile column shows the results for those students in principles of microeconomics or macroeconomics courses at the beginning of the course. The posttest percentile column shows the results from the same group of students at the end of the course.

Percentile ranks allow comparisons to be made among groups or individuals. In this respect Tables 4 and 5 are particularly valuable and more useful than the aggregate statistics reported in Table 3. If it is determined that the overall specifications on the TUCE-4 are appropriate for a particular principles class, the data in Table 4 can be interpreted in the following way: A class with a micro pretest mean score of 12 and a posttest mean score of 17 would be in the 84th percentile of individual students in both cases, whereas a class with a pretest mean score of 10 and posttest mean score of 15 would be in the 67nd percentile on the pretest and the 74th percentile on the posttest. Thus, a pre- to posttest gain in mean scores

TABLE 4. Distribution of Pre- and Posttest Scores on Micro TUCE-4: *Matched*

	Pretest (n = 3255)			**Posttest** (n = 3255)		
Raw Score	**No. of Scores**	**Percentile Rank**	**T-Score**	**No. of Scores**	**Percentile Rank**	**T-Score**
30				1		87
29				1		85
28				4		83
27				4		80
26				11		78
25	1		97	25	99	76
24	2		94	18	99	74
23	4		91	38	98	72
22	2		88	47	97	70
21	10		85	69	95	68
20	9	99	82	89	93	65
19	15	99	79	99	91	63
18	18	99	76	118	88	61
17	36	98	73	167	84	59
16	43	97	70	170	79	57
15	83	96	67	207	74	55
14	127	93	64	229	67	53
13	163	89	61	273	60	50
12	248	84	58	260	52	48
11	320	77	55	288	44	46
10	340	67	52	283	35	44
9	451	56	49	262	26	42
8	426	43	46	203	18	40
7	349	29	43	166	12	38
6	287	19	40	105	7	36
5	176	10	37	63	4	33
4	88	5	34	32	2	31
3	32	2	31	16	1	29
2	17	1	28	5	0	27
1	8	0	25	2	0	25
Mean Score			9.39			12.77
Std. Deviation			3.32			4.68
Alpha			.46			.70
Std. Error of Measurement			2.45			2.58

TABLE 5. Distribution of Pre- and Posttest Scores on Macro TUCE-4: *Matched*

	Pretest (n = 2789)			**Posttest** (n = 2789)		
Raw Score	**No. of Scores**	**Percentile Rank**	**T-Score**	**No. of Scores**	**Percentile Rank**	**T-Score**
30				3		80
29				9		78
28				12		76
27	1		99	11	99	74
26	0		97	26	99	72
25	0		94	40	98	70
24	3		91	45	96	69
23	4		88	68	95	67
22	5		85	76	92	65
21	9		82	109	90	63
20	7	99	79	102	86	61
19	18	99	76	105	82	59
18	31	98	74	118	78	57
17	40	97	71	141	74	55
16	53	96	68	173	69	53
15	95	94	65	192	63	52
14	137	91	62	191	56	50
13	161	86	59	205	49	48
12	203	80	56	217	42	46
11	270	73	53	209	34	44
10	322	63	51	171	27	42
9	363	51	48	172	20	40
8	323	38	45	150	14	38
7	284	27	42	89	9	36
6	230	17	39	81	6	34
5	127	8	36	40	3	33
4	69	4	33	23	1	31
3	20	1	30	8	0	29
2	13	1	28	1	0	27
1	1	0	25	2	0	25
Mean Score			9.80			14.19
Std. Deviation			3.48			5.29
Alpha			.51			.77
Std. Error of Measurement			2.45			2.53

of five questions would be interpreted differently in these two situations. Likewise, Table 5 indicates that a class with a pretest mean of 10 and a posttest mean of 18 on the macro TUCE-4 would be in the 63rd percentile of individual students on the pretest and the 78th percentile on the posttest. This would indicate that the average performance of students in this class had increased relative to the national norms on the TUCE-4. Alternatively, a pretest mean of 10 on the macro TUCE-4 (63th percentile) and a posttest mean of 15 (63th percentile) would indicate that the average performance of the students had stayed constant relative to the national norms.

T-scores are also reported in Tables 4 and 5. T-scores transform the raw scores to a scale with a mean of 50 and a standard deviation of 10. This transformation makes it easier to determine how many standard deviations (measured by 10 units on a T-score scale) a test score is above or below the mean of 50. To do the linear conversion, the *z*-score for each raw test score is calculated by taking the difference between a test score and the mean and dividing it by the standard deviation for the test. The *z*-score value is then multiplied times 10 and this value is added to 50 to create the T-score. For example, assume that on the micro test an individual student or class has a raw mean score of 22. The *z*-score for this raw score is 1.97 [(22-12.77)/4.68 = 1.97]. This *z*-score is then multiplied times 10 and added to 50, and rounded, to produce a T-score of 70. This T-score of 70 indicates that a raw score of 22 is 2 standard deviations above the mean.

Item Analysis

Tables 6 and 7 report the individual item results for the questions on each test. Economics instructors will want to know how their students performed on certain items of the TUCE-4. Information on item difficulty and discrimination will be particularly important in cases where the instructor covered only some of the concepts or topics included in the test, and will help instructors evaluate student performance.

Item Difficulty. Table 6 shows the percentage of correct responses on each item for students taking the micro TUCE-4 as a pretest and posttest. Table 7 reports the same item information for students taking the macro TUCE-4 as a pretest and posttest. This percentage is an estimate of the difficulty of an item for a particular group of students. Theoretically, this percentage can range from 0 to 100 percent, but most items will fall in the 30 to 70 percent correct range for those students taking the TUCE-4 as a posttest. Students taking the TUCE-4 as a pretest will generally have a lower percentage correct for each item because the students will not yet have been taught economics.

Data on item difficulty should be interpreted with care. Item difficulty (percentage of correct responses) depends on many things besides the complexity of the fact, concept, or principle being tested. Such matters as course emphasis on the item content question, the closeness or plausibility of incorrect alternatives or "distractors" and the relation of the item content to such factors as the students' previous education, work experience, and reading may also affect item difficulty. It is worth emphasizing, therefore, that undue attention should not be placed on small differences between the percentages reported in this manual and those obtained in the classroom.

Each question on the TUCE-4 has four possible choices: one correct answer and three distractors. Pure chance would dictate an expected correct score of 25 percent on the test for those who had no knowledge of economics. If some students with economics instruction score below 25 percent on the test (or about 7 or less correct answers), their answer sheets in particular should be carefully checked for systematic errors in test marking, scoring, or test administration.

Item Discrimination. Also reported in Tables 6 and 7 is the discrimination coefficient for each TUCE-4 micro or macro item. This coefficient is the corrected item-to-total score correlation or point-biserial correlation (R_t). The coefficient measures the correlation between the students' total test score, adjusted or corrected by

TABLE 6. Item Analysis: TUCE-4 Micro Pre-Post: *Matched* (n = 3255)

Item	Correct Answer	Corrected Item—Total Correlation	*Percent Correct* Posttest	Pretest
1	A	.26	50%	39%
2	B	.18	40	33
3	A	.24	50	36
4	A	.22	57	14
5	C	.17	46	40
6	C	.31	46	23
7	D	.30	49	45
8	A	.34	37	21
9	D	.19	31	22
10	A	.11	44	37
11	A	.33	32	11
12	C	.23	45	24
13	B	.17	50	37
14	B	.21	45	30
15	C	.18	34	22
16	C	.23	50	43
17	D	.33	43	32
18	B	.20	41	30
19	C	.15	43	43
20	C	.19	31	17
21	D	.23	45	43
22	A	.17	59	56
23	C	.22	31	24
24	B	.20	49	41
25	D	.28	34	23
26	D	.12	34	29
27	B	.15	41	29
28	B	.20	35	24
29	A	.26	37	31
30	D	.29	49	41

TABLE 7. Item Analysis: TUCE-4 Macro Pre-Post: *Matched* (n = 2789)

Item	Correct Answer	Corrected Item—Total Correlation	*Percent Correct* Posttest	Pretest
1	A	.31	53%	23%
2	B	.21	61	49
3	C	.31	69	46
4	D	.33	46	36
5	A	.31	59	11
6	B	.26	47	33
7	B	.25	60	51
8	C	.24	50	41
9	C	.26	33	22
10	B	.19	41	35
11	D	.41	59	34
12	B	.17	55	40
13	B	.30	63	56
14	A	.33	48	25
15	B	.26	61	50
16	C	.17	38	28
17	C	.26	37	31
18	A	.31	45	17
19	A	.22	40	33
20	C	.33	60	51
21	A	.29	42	20
22	D	.34	33	18
23	A	.30	36	26
24	D	.23	33	26
25	B	.36	60	44
26	C	.22	31	22
27	A	.24	33	17
28	D	.32	51	35
29	D	.31	34	25
30	D	.32	44	34

omitting that particular item, and the dichotomous (right or wrong) score on the particular item.[2]

The correlation provides an assessment of the functioning of that item with the students who were tested. This correlation coefficient ranges from 0 to 1. The *higher* the value of the coefficient, the better the item functions as a discriminator between those students who know more or know less economics. A coefficient of zero indicates an item fails to discriminate between those with more and less knowledge of economics as measured by their total score. Questions with a *negative* coefficient are *reverse discriminators* (indicating that more lower-scoring students get the question right than do higher-scoring students). In general, if an item has a discrimination coefficient below 0.10, the item may either be a weak discriminator or it may indicate that there was little or no instruction on the tested concept.

Item Responses. Tables 8 and 9 show the percentages of students selecting one of four options or leaving the answer blank on the micro or macro TUCE-4. These data are supplied for both the pretest and the posttest. The percentage for the correct response is shown in boldface and with an asterisk.

An analysis of item responses can be useful for test users. Table 8 shows detailed pretest and posttest item data for all 30 micro items. Test users will need to have a copy of the test at hand (Appendix 4) to see the questions to which these data relate. For example, if a substantial percentage of students answered A when the correct answer was C, the instructor would do well to study distractor A to determine the reason why students selected the incorrect response. Table 9 shows detailed pretest and posttest item data for all 30 macro questions. Test users will again need to have a copy of the test at hand (Appendix 5) to see the questions to which these data relate.

[2]The formula for a point biserial correlation between an individual test item, g, and the total test score (R_t) is:

$$R_{\bar{X}g} = \left[\bar{X}_g - \bar{X}/S_X\right]\sqrt{P_g/Q_g}$$

where $\bar{X}_g$ = mean score of those answering item g correctly; $\bar{X}$ = mean score of the total test; S_x = standard deviation on the total test; P_g = proportion answering item g correctly; $Q_g = 1 - P_g$.

Item Examples and Interpretation. Three sample questions are shown below to illustrate the content and cognitive classifications of test of items in Tables 1 and 2, to indicate how test items were constructed, and to demonstrate how the item analysis data presented in Tables 6–9 can be interpreted. The data following each sample question show the percentage of students in the pretest norming samples selecting each alternative before (pre) and after (post) taking a principles of economics course. Also shown with each question is the corrected item-to-total correlation or point-biserial correlation (R_t).

Macro Question #4. Content Category "B." Cognitive Category "RU."

The limit of total productive capacity in an economy is set by:

A. the amount of money in circulation.
B. business demand for goods and services.
C. the amount of government spending and taxation.
D. the quantity and quality of its productive resources.

Pre	*Post*	
13%	12%	
42	32	
7	9	
36	**46**	**R_t=.33**

Micro Question #3. Content Categories "B." Cognitive Category "EA."

If all of the firms in a competitive industry are legally required to meet new regulations that increase their costs of production:

A. supply of the product will decrease.
B. demand for the product will decrease.
C. the long-run economic profits of the individual firms in the industry will decrease.
D. the short-run economic profits of the individual firms in the industry will decrease.

TABLE 8. Percentage Response to Items: TUCE-4 Micro Pre- and Posttests: *Matched* (*n = 3255*)

Item	A	B	C	D	Blank
1Post	**50***	33	14	3	0.1
1Pre	**39***	45	13	3	0.5
2	5	**40***	23	31	0.5
2	6	**33***	23	37	1.4
3	**50***	10	28	11	0.6
3	**36***	15	36	12	0.9
4	**57***	18	19	6	0.6
4	**14***	29	35	19	1.9
5	32	9	**46***	13	0.9
5	29	11	**40***	19	1.3
6	8	9	**46***	36	0.5
6	13	12	**23***	51	0.9
7	7	16	28	**49***	0.5
7	7	19	29	**45***	1.0
8	**37***	5	3	55	1.1
8	**21***	3	1	74	1.4
9	33	11	26	**31***	0.5
9	33	10	35	**22***	0.8
10	**44***	12	7	36	0.5
10	**37***	10	10	43	1.1
11	**32***	17	22	29	0.7
11	**11***	24	16	48	1.6
12	17	15	**45***	23	0.9
12	13	24	**24***	37	2.0
13	19	**50***	21	10	0.4
13	7	**37***	38	17	1.0
14	18	**45***	14	22	1.0
14	24	**31***	14	30	1.8
15	24	11	**34***	31	0.6
15	27	12	**22***	38	1.2
16	8	29	**50***	11	1.3
16	7	33	**43***	15	1.9
17	12	27	16	**43***	1.6
17	13	26	27	**32***	2.0
18	25	**41***	28	5	0.7
18	30	**30***	31	8	1.0
19	10	13	**44***	32	1.0
19	15	14	**43***	27	1.0
20	11	37	**31***	21	1.4
20	16	49	**17***	17	1.8
21	17	15	22	**45***	1.5
21	14	16	25	**43***	2.6
22	**59***	25	9	6	1.5
22	**56***	19	14	9	2.0
23	18	27	**31***	21	2.4
23	14	36	**24***	24	2.3
24	22	**49***	13	14	2.4
24	26	**41***	16	16	1.9
25	14	23	27	**34***	2.3
25	16	26	32	**23***	2.8
26	12	45	6	**34***	2.9
26	16	42	9	**29***	3.5
27	10	**41***	18	28	2.8
27	16	**29***	21	31	3.1
28	20	**35***	37	6	2.7
28	15	**24***	50	8	2.9
29	**37***	11	40	9	2.8
29	**31***	16	40	11	3.0
30	15	13	19	**49***	4.0
30	17	17	22	**41***	3.7

Note: *Correct answer

TABLE 9. Percentage Response to Items: TUCE-4 Macro Pre- and Posttests: *Matched* (*n = 2789*)

Item	A	B	C	D	Blank
1Post	**53***	5	9	33	0.5
1Pre	**23***	5	13	59	0.6
2	20	**61***	13	5	0.7
2	27	**49***	15	8	0.6
3	2	12	**69***	16	0.5
3	2	19	**46***	33	0.7
4	12	32	9	**46***	1.1
4	13	43	7	**36***	0.8
5	**59***	6	25	10	0.6
5	**12***	18	44	25	0.7
6	22	**47***	15	14	1.2
6	26	**33***	22	18	1.0
7	16	**60***	16	7	0.6
7	19	**51***	23	7	0.3
8	22	8	**50***	19	0.6
8	22	11	**41***	25	0.6
9	17	30	**33***	20	0.7
9	19	35	**22***	23	0.9
10	28	**41***	21	10	0.6
10	32	**35***	17	16	0.7
11	12	17	11	**59***	1.0
11	18	34	14	**34***	0.7
12	5	**55***	11	28	0.9
12	5	**40***	18	37	0.6
13	14	**63***	13	9	0.6
13	11	**57***	18	14	0.5
14	**48***	15	20	16	0.9
14	**25***	23	28	23	0.8
15	6	**61***	11	21	0.7
15	11	**50***	24	16	0.3
16	20	24	**38***	18	0.8
16	26	28	**28***	18	0.6
17	35	18	**37***	9	0.9
17	39	17	**31***	13	0.4
18	**45***	10	26	18	0.8
18	**17***	19	47	17	0.5
19	**40***	26	19	15	1.0
19	**33***	28	24	14	1.0
20	12	21	**60***	6	1.0
20	17	23	**51***	8	0.6
21	**42***	8	18	31	1.5
21	**20***	8	35	36	0.9
22	24	30	11	**33***	1.6
22	31	36	14	**18***	1.1
23	**36***	26	24	13	1.3
23	**26***	32	24	16	1.1
24	15	27	23	**33***	1.4
24	16	32	25	**26***	1.3
25	11	**60***	14	13	1.3
25	18	**44***	22	16	0.9
26	22	15	**31***	32	1.3
26	27	16	**22***	33	1.2
27	**33***	22	24	20	1.5
27	**17***	22	28	31	1.3
28	17	17	13	**51***	1.5
28	21	24	19	**35***	1.5
29	19	16	30	**34***	1.7
29	20	19	35	**25***	1.5
30	13	23	19	**44***	1.7
30	17	26	21	**34***	1.7

Note: *Correct answer

Pre	*Post*	
36%	**50%**	**R_t=.24**
15	10	
36	28	
12	11	

Micro Question #16. Content Category "E." Cognitive Category "IA."

"The effect of an excise tax on the products of pollution-producing industries will be a cutback in production. If the tax was levied directly on the amount of pollution generated, the long-run cutbacks in production would be much smaller." This statement is most likely to be:

A. false, provided the amount of the taxes on products and pollution is equal.
B. false, because most firms would rather pay the tax than cut back production.
C. true, because firms would have a greater incentive to adopt new technology that causes less pollution.
D. true, because most taxes levied on pollution affect the demand curve; taxes on product affect the supply curve.

Pre	*Post*	
7%	8%	
33	29	
43	**50**	**R_t=.23**
15	11	

All three of these items illustrate the point that, unless there was a content or format reason for doing otherwise, the alternatives on each question are arranged uniformly from the shortest to the longest. Similarly, special care was taken to ensure that each of the alternatives (A, B, C, or D) is the correct option about the same number of times. These changes were made so that the longest alternative, which some "testwise" students may think is usually the correct alternative, does not call attention to itself, and to give no advantage in selecting or guessing answers based on its placement in the set. The TUCE-4 score should be a measure of economic understanding rather than multiple-choice test-taking skills.

All three of the sample questions have reasonably good statistical properties. All the alternatives were plausible and attracted some student response. Also, the point biserial correlation (R_t) between the mean score of students choosing the correct answer and the mean total test score is positive and shows that the items do discriminate between those students with more or less economic understanding.

Reliability

The reliability of a test is the degree of consistency with which a test measures student performance. For example, two students taking the same test are likely to obtain different scores, but each student taking the test again (without intervening instruction in economics) should obtain about the same score as the first time. Many factors (including practice in taking the test or guessing) cause changes in student performance from day to day. As a result, we can never measure a student's performance perfectly (that is, obtain a student's "true" score).

Alpha. One measure of overall test reliability is the coefficient alpha (Cronbach, 1951).[3] It is a measure of the internal consistency among test items with a common focus, which for the TUCE-4 is either microeconomics or macroeconomics. One way to conceptualize internal consistency is to think of splitting the test in half and correlating student scores on both halves. The alpha coefficient provides an estimate of the average of all possible split half correlations.

The alpha statistic ranges from zero to 1.00. The higher the coefficient, the better items work together in measuring the test construct, and thus the greater the statistical reliability of the test. An alpha of 1.00 would indicate a perfectly reliable test, while a coefficient of zero would indicate a totally unreliable one.

[3]The formula for the Cronbach alpha is

$$\alpha = n / (n-1)\left[1 - (\sum V_i / V_t)\right]$$

where n = number of test items; V_t = variance of the total test; and $\sum V_i$ = sum of the variance of individual items.

As shown in Table 3, the alpha is 0.70 for the micro TUCE-4 and .77 for the macro TUCE-4. These alphas indicate that there is reasonably good internal consistency among items and that each test is a reliable measure of achievement in items covering principles of microeconomics or macroeconomics. The alpha for the TUCE-4 macro test is the same as the alpha report for the 30-item version of TUCE-3. The alpha for the micro test is lower than the .80 reported for TUCE-3, probably reflecting the fact that the TUCE-4 exam is a more difficult test for students, and perhaps the somewhat expanded range of topics and techniques covered on this version of the exam (international, game theory, graphical analysis, etc.). Regardless of the alpha estimate of reliability, the major question to be determined by each user of the TUCE-4 is whether the test as a whole (or individual questions) is appropriate for the testing of his or her students.

SEM. It is also possible to estimate the amount of variation in test scores that is due to measurement error, and therefore to specify a range within which one can be relatively certain the "true" score will fall. By taking account of such measurement error, the reliability of the test as a whole can be estimated.

The standard error of measurement (SEM), which is reported in Table 3, is an estimate of the amount of variation that can be expected in a test score (Linn & Gronlund, 2000, pp. 119-125).[4] The standard error of measurement for posttest scores on the TUCE-4 is 2.58 for micro and 2.53 for macro. For approximately two thirds of the scores obtained, the error of measurement will be 2.58 or less for micro and 2.53 or less for macro. For about 95 percent of the scores obtained, the error measurement will be two SEMs or less. The smaller the SEM, the more accurate a test is as a measuring instrument for student achievement. Individual test scores are best thought of as lying within a range, rather than as a single score, because of our inability to measure perfectly (the SEM is never zero).

Validity

Substantial evidence was collected for establishing the validity of the TUCE-4 as an achievement measure of understanding of principles of microeconomics and macroeconomics. This evidence reported in this section consists of content and construct validity.

Content. One of the most important validity questions for an achievement test such as the TUCE-4 is whether or not it measures what *ought* to be measured. This question cannot be answered by reference to statistics. The work that was done to establish the *content validity* of the TUCE-4 was described in detail in Sections 1 and 2 of this manual. The results of this content validity work are shown in the content specification tables (Tables 1 and 2).

The process used for test development also ensured that the TUCE-4 items would contain valid content considered to be important for teaching principles of microeconomics and macroeconomics. The initial group of test items came from previous editions of the TUCE, which had undergone review by several committees of economics professors and had been tested with students for those editions. These initial items were revised as necessary and new test items were written to cover additional content or to replace previous items with content problems. The content of all items on the TUCE-4 were reviewed by the test development committee that was composed largely of professors teaching micro or macro principles of economics. The test items were then field-tested with students taking micro or macro principles to see how they worked, and then the item content was reviewed again by the test developers. In addition, the content of all test items were reviewed by a National Advisory Committee that was composed of distinguished economists and changes were made based on their recommendations.

[4]The standard error of measurement (SEM) is the standard deviation multiplied by $\sqrt{1-r}$, where r is the estimate of reliability. SEM can be used to interpret individual test scores in a manner similar to that for the standard deviation. For example, for about two-thirds of the students taking a test, the "true score" will not deviate more than ± SEM from the score they actually obtain.

The content validity of the TUCE-4 was determined by comparing the test questions with the content specification determined to be important by authoritative experts in economics. Nevertheless, there is no one standard for content validity. Whether the TUCE-4 is a valid test of principles of microeconomics or macroeconomics often depends on the purpose for which it is used. Some economics instructors or test users may disagree with the economics on the TUCE-4. In those cases, the TUCE-4 may not be content valid for the purposes for which the test users want to use the test. For most economics instructors, however, the TUCE-4 should be a useable standardized test for measuring student achievement in a principles course.

Construct. There is substantial evidence from the student sample on the *construct* validity of the TUCE-4. Construct validity refers to the ability of the test to measure the underlying construct or focus of the test, which for this test is principles of microeconomics or macroeconomics. One type of evidence for construct validity that is presented is whether students score substantially better on the posttest than the pretest.

As shown in Table 3, microeconomics students scored 3.4 points higher on the TUCE-4 posttest than on the pretest (a 36 percent improvement over the pretest score). Macroeconomics students scored 4.4 points higher on the posttest than the pretest (a 45 percent improvement over the posttest score). Both score differences are statistically significant in the expected direction. The probability that this difference is due to chance is about zero (probability less than 0.001).

A further check on the construct validity of any individual test item may be made by reviewing the performance on each item for students on the posttest and the pretest (Tables 6 and 7). By comparing the percent correct from each group, it is clear that the posttest item percent correct is higher than the pretest item percent correct on all items except one. The exception is #19 on the micro test where the proportion stays the same, perhaps because the content was not taught.

More Construct Validity Evidence. Tables 10 and 11 present some descriptive statistics from the matched student samples for the micro and macro TUCE-4. The mean TUCE-4 scores, standard deviations, and sample sizes are given for each subgroup on the pretest and posttest. With the exception of institution type, the data were all self-reported by students and collected from a student survey completed at the time of the pretest and/or posttest.

The data are broken down by some fifteen factors that cover a wide range of student characteristics: (1) gender; (2) age; (3) year in school; (4) race or ethnicity; (5) communicate better in English than another language; (6) communicate equally well in English and another language; (7) type of institution attended; (8) enrollment status; (9) grade point average; (10) academic major; (11) number of economics courses taken; (12) plans to take more economics courses; (13) expected grade in the economics course at the time of the pretest; (14) expected grade in the economics course at the time of the posttest; and, (15) the number of calculus courses taken.

Tables 10 and 11 show higher posttest scores compared with pretest scores across *all* fifteen student characteristics that are listed and within each category of each characteristic. What these results indicate is that performance on the test is responsive to economics instruction, regardless of other personal characteristics of the student. The significant increases in scores from the pretest to the posttest across all these different characteristics of students indicate that there is *construct* validity to the TUCE-4. The results suggest that differences from pretest to posttest are most likely due to economic instruction and are not likely attributable to some other factor or characteristic associated with each student.

It is important to stress that these categorical breakdowns must be interpreted with caution, because some of the cell sizes (the subgroup *n*) are small. The breakdowns are also for single characteristics without controlling for other characteristics.

TABLE 10. Descriptive Statistics for Groups within the *Matched* Norming Sample: TUCE-4 Micro

	Pretest			Posttest		
	Mean	**Std. Dev.**	**Number**	**Mean**	**Std. Dev.**	**Number**
Gender						
Females	9.04	3.14	1,384	12.31	4.44	1,384
Males	9.63	3.42	1,848	13.12	4.83	1,848
Age						
18 or younger	9.40	3.18	516	14.02	4.98	516
19	9.35	3.21	866	12.78	4.63	866
20	9.35	3.19	732	12.21	4.49	732
21	9.38	3.37	421	12.58	4.67	421
22 or older	9.49	3.67	686	12.57	4.56	686
Year in school						
Freshman	9.40	3.19	771	13.96	4.91	771
Sophomore	9.40	3.27	1,175	12.47	4.59	1,175
Junior	9.37	3.32	932	12.06	4.39	932
Senior	9.37	3.65	302	12.98	4.52	302
Other	9.39	4.19	61	13.72	5.66	61
Race/Ethnic Origin						
White	9.54	3.41	2,204	13.22	4.76	2,204
African-American/Black	8.33	2.91	314	11.21	4.02	314
Hispanic/Latino	9.36	3.03	170	11.94	4.25	170
American Indian/Alaskan Native	9.77	3.02	30	12.83	5.13	30
Other	9.32	3.14	501	12.06	4.52	501
Communicate better in English than another language						
Yes	9.32	3.30	2,743	12.63	4.65	2,743
No	9.80	3.45	475	13.61	4.75	475
Communicate equally well in English and another Language						
Yes	9.18	3.42	468	11.89	4.32	468
No	9.43	3.14	2,706	12.95	4.72	2,706
Institution type						
Associate's College	9.35	3.11	252	11.21	4.12	252
Baccalaureate College	9.67	3.50	273	15.18	5.02	273
Master's College / University	9.22	3.23	1,972	12.38	4.64	1,972
Doctoral/Research University	9.73	3.50	758	13.44	4.45	758
Enrollment Status						
Full-time	9.35	3.29	2,980	12.78	4.68	2,980
Part-time	9.87	3.57	232	12.93	4.69	232

TABLE 10. Descriptive Statistics for Groups within the *Matched* Norming Sample: TUCE-4 Micro (continued)

	Pretest			Posttest		
	Mean	**Std. Dev.**	**Number**	**Mean**	**Std. Dev.**	**Number**
GPA						
None	9.49	2.85	243	15.02	4.83	243
< 1.99	8.61	3.32	41	11.24	4.41	41
2.00–2.99	8.81	3.05	1,198	11.35	4.13	1,198
3.00–4.00	9.66	3.41	1,630	13.46	4.74	1,630
Major						
Economics	10.50	3.93	222	14.10	5.31	222
Business	9.07	3.14	1,636	12.24	4.37	1,636
Other	9.61	3.34	1,325	13.24	4.88	1,325
Number of Economics Courses Taken						
None	9.17	3.15	1,958	13.00	4.71	1,958
One	9.68	3.51	1,107	12.49	4.61	1,107
Two	9.81	3.75	139	12.07	4.89	139
Three	9.71	3.69	17	11.71	4.28	17
Four or more	12.46	2.79	13	13.54	4.31	13
Plan to Take More Economics Courses						
Yes, to meet requirements	9.32	3.26	1,491	12.92	4.56	1,491
Yes, if fits into my schedule	10.01	3.56	344	13.86	5.42	344
No	9.31	3.31	1,125	12.37	4.62	1,125
Expected Grade for Course at Pretest						
A	9.82	3.51	1,831	13.02	4.84	1,831
B	8.81	2.94	1,262	11.82	4.25	1,262
C	8.30	2.71	79	11.21	3.87	79
Expected Grade for Course at Posttest						
A	10.83	3.79	751	15.25	5.41	751
B	9.35	3.16	1,215	12.78	4.17	1,215
C	8.42	2.72	878	11.21	3.87	878
D	7.87	2.50	112	9.80	3.63	112
Number of Calculus Courses Taken						
None	9.13	3.17	2,089	12.41	4.46	2,089
One	9.65	3.29	777	13.16	4.78	777
Two	10.09	3.83	243	13.76	5.21	243
Three	10.73	4.18	123	14.81	5.71	123

Note: All data are self-reported from a student survey except for institutional type.

TABLE 11. Descriptive Statistics for Groups within the *Matched* Norming Sample: TUCE-4 Macro

	Pretest			Posttest		
	Mean	**Std. Dev.**	**Number**	**Mean**	**Std. Dev.**	**Number**
Gender						
Females	9.24	3.26	1,124	13.37	5.02	1,124
Males	10.20	3.57	1,651	14.77	5.39	1,651
Age						
18 or younger	9.93	3.49	290	14.00	4.84	290
19	9.56	3.34	773	13.61	5.13	773
20	9.55	3.30	693	14.03	5.20	693
21	9.53	3.44	395	14.21	5.27	395
22 or older	10.50	3.77	610	15.28	5.66	610
Year in school						
Freshman	9.76	3.49	480	14.19	4.77	480
Sophomore	9.68	3.44	1,238	13.83	5.28	1,238
Junior	9.73	3.33	761	14.51	5.45	761
Senior	10.16	3.82	227	14.64	5.30	227
Other	11.63	4.01	79	15.81	6.32	79
Race/Ethnic Origin						
White	10.00	3.49	1,974	14.76	5.22	1,974
African-American/Black	8.54	2.62	254	11.20	4.47	254
Hispanic/Latino	8.80	3.41	123	11.44	4.75	123
American Indian/Alaskan Native	8.89	2.49	47	14.60	5.38	47
Other	10.10	3.80	380	14.17	5.37	380
Communicate better in English than another language						
Yes	9.78	3.43	2,440	14.14	5.26	2,440
No	9.97	3.86	335	14.67	5.47	335
Communicate equally well in English and another Language						
Yes	9.65	3.90	333	13.42	5.45	333
No	9.82	3.41	2,407	14.32	5.263	2,407
Institution type						
Associate's College	9.59	3.40	186	13.10	4.86	186
Baccalaureate College	9.04	3.11	383	13.15	4.62	383
Master's College / University	9.86	3.34	1,464	14.32	5.04	1,464
Doctoral/Research University	10.13	3.87	757	14.76	6.03	757
Enrollment Status						
Full-time	9.77	3.46	2,572	14.16	5.26	2,572
Part-time	10.33	3.85	196	15.05	5.51	196

TABLE 11. Descriptive Statistics for Groups within the *Matched* Norming Sample: TUCE-4 Macro (continued)

	Pretest			Posttest		
	Mean	**Std. Dev.**	**Number**	**Mean**	**Std. Dev.**	**Number**
GPA						
None	10.20	3.35	130	13.59	4.66	130
< 1.99	9.05	3.57	56	11.68	5.06	56
2.00–2.99	9.02	3.05	1,050	12.48	4.60	1,050
3.00–4.00	10.08	3.48	1,333	14.91	5.26	1,333
Major						
Economics	10.72	4.48	128	15.05	6.32	128
Business	9.53	3.24	1,453	13.92	5.04	1,453
Other	10.04	3.60	1,180	14.46	5.46	1,180
Number of Economics Courses Taken						
None	9.47	3.31	1,836	13.58	5.10	1,836
One	10.36	3.62	821	15.34	5.35	821
Two	10.78	3.81	97	15.53	5.94	97
Three	11.67	5.17	18	16.50	7.04	18
Four or more	13.09	5.03	11	17.64	4.84	11
Plan to Take More Economics Courses						
Yes, to meet requirements	9.71	3.43	1,474	13.95	5.13	1,474
Yes, if fits into my schedule	10.98	3.82	320	16.08	5.66	320
No	9.61	3.40	897	14.20	5.29	897
Expected Grade for Course at Pretest						
A	10.17	3.67	1,571	15.05	5.54	1,571
B	9.29	3.13	1,065	13.04	4.70	1,065
C	9.20	2.98	83	12.29	4.14	83
Expected Grade for Course at Posttest						
A	11.28	3.98	699	17.20	5.72	699
B	9.67	3.24	1,096	14.14	4.87	1,096
C	8.90	3.04	793	12.27	4.24	793
D	8.56	2.59	89	10.37	3.54	89
Number of Calculus Courses Taken						
None	9.39	3.22	1,737	13.31	4.91	1,737
One	10.00	3.42	619	14.64	5.23	619
Two	11.21	4.20	231	16.90	5.71	231
Three	11.23	4.12	193	17.49	5.63	193

Note: All data are self-reported from a student survey except for institutional type.

5. USING THE TUCE-4

There are many possible uses for the TUCE-4. These uses can include group or individual evaluation of student achievement, as a teaching strategy, or for a research study. Before describing these possible uses one note of caution about test security is necessary.

Test Security

Instructors using the TUCE-4 for testing, instruction, or research should be aware of the potential test security problems that may arise when the same questions are used over time. Any widely available test poses certain security problems that need to be considered and checked before scores are used for grading or research purposes. To help maintain the reliability and validity of the TUCE-4, test users should not do a public release of test items to students.

Group Evaluation

By comparing the scores of their students with the percentile distributions in Tables 4 and 5, economics instructors can estimate the comparative effectiveness of courses in achieving the objectives measured by the TUCE-4. The concept of reliable group performance (as reflected in the mean score) is different from the concept of reliable individual performance. Group means can be reliable even when scores for individuals may not be, and it is possible to get reliable group mean scores on the TUCE-4 with relatively small groups of students.

In addition to the mean scores of groups, instructors may also be interested in the percent of students answering individual questions correctly. The norming data in tables 6 to 9 permit such comparisons for each question on the TUCE-4. And, as indicated above, an *approximate* norm group mean for a given set of questions can be obtained by aggregating the data showing the percent correct for each question in the set.

Before judging the adequacy of his or her students' performance in comparison to the norming data published in the *TUCE-4 Manual*, a test user should examine the TUCE-4 in relation to the content and purposes of the courses taught and the characteristics of the students in comparison to those in the national norming samples. The test items on each form of the TUCE-4 are only samples—albeit carefully selected samples—of the possible questions that might have been used; and they are weighted by the specifications determined by the TUCE-4 committee. Each test user must independently decide the extent to which the emphasis of his or her courses agrees with that of the tests.

Individual Student Evaluation

Although the TUCE-4 was not primarily designed for this purpose, economics instructors may wish to use the tests to help evaluate the achievement of individual students. Test validity and reliability are especially important when the TUCE-4 is used in this way. The primary question regarding validity is whether the TUCE-4 and the economics instruction provided in a course to students are congruent. Congruence cannot be measured statistically; it is a matter of judgment by the instructor. A test that is not properly related to a course of study is not a valid measure of achievement in that course. Before using the TUCE-4 for evaluating students, instructors should analyze each test question in order to judge whether or not the test is valid for their particular courses. The questions on the TUCE-4, of course, can also be supplemented with other questions to ensure that the total "package" of evaluation questions is congruent.

Any measure of student achievement contains a margin of error that can be estimated by the standard error of measurement (SEM) of a test score (Table 3 and Note 4). This statistic indicates the amount of variation that may be expected in a test score. For example, a raw score of 15 on the micro test with an SEM of 2.58 indicates about 67 percent certainty that a person's

"true" score lies in a range from 12.42 to 17.58 (15 +/– 2.58), or that we can be 95 percent certain that the "true" score lies in a range from 9.84 to 20.16 [15 +/– (2 x 2.58)].

Teaching Aid

Many instructors will find the TUCE-4 useful as a basis for teaching. The development of the tests according to careful specifications has resulted in coverage of a broad range of economic knowledge in each test. The TUCE-4 can be administered as a pretest to help in planning the course of instruction. If student pretest scores are high or low on particular topics, the instructor may wish to use this information for planning content coverage or assignments. The TUCE-4 can be administered as a posttest to help students check their understanding. In this case, the instructor can discuss with students the reasons for the correct or incorrect answers to test items. This practice, however, should be done in class and under secure conditions so test items do not get circulated to other students, thus invalidating the test for future use.

Research Studies

Researchers in economic education are often interested in different problems than are instructors who wish to compare their classes with the national norms. These studies require a greater sophistication in obtaining and interpreting data to answer a research question. Past editions of the TUCE were used in many research studies and have been discussed in several reviews of the research literature and research methods (Siegfried and Fels, 1979; Becker, 1983a, 1983b, and 1983c; Becker and Walstad, 1987, Siegfried and Walstad, 1998; and Becker, 1997). The *Journal of Economic Education* is also another key resource with research findings for those test users interested in conducting research studies. The TUCE-4 can be used for research studies that seek to compare student achievement in principles across different institutions or classes using a standardized measure. The availability of this instrument should help advance research in economic education, just as with the three previous editions of this test.

REFERENCES

Becker, William E. 1997. "Teaching economics to undergraduates," *Journal of Economic Literature, 35* (3) (September), 1347-1373.

_____. 1983a. "Economic education research: Part I, issues and questions," *Journal of Economic Education, 14* (Winter), 10-17.

_____. 1983b. "Economic education research: Part II, new directions in theoretical model building," *Journal of Economic Education, 14* (Spring), 4-10.

_____. 1983c. "Economic education research: Part III, statistical estimation methods," *Journal of Economic Education, 14* (Summer), 4-15.

_____, and Walstad, William B., eds. 1987. *Econometric modeling in economic education research.* Boston: Kluwer-Nijhoff.

Bloom, Benjamin S., ed. 1956. *Taxonomy of educational objectives*. New York: David McKay.

Cronbach, L.J. 1951. "Coefficient alpha and the internal structure of tests," *Psychometrica*, *16*, 297-334.

Fels, Rendigs. 1967. "A new test of understanding in college economics," *American Economic Review, Papers and Proceedings, 57* (2), 660-666.

Linn, R.L., and Gronlund, N.E. 2000. *Measurement and assessment in teaching* (8th ed.). Upper Saddle River, NJ: Prentice-Hall.

Saunders, Phillip. 1981. *Revised test of understanding in college economics: Interpretive manual*. New York: National Council on Economic Education.

_____. 1991a. *Test of understanding of college economics: Examiner's manual* (Third Edition). New York: National Council on Economic Education.

_____ 1991b. "The third edition of the test of understanding of college economics," *American Economic Review, Papers and Proceedings, 81* (2), 32-37.

_____, Fels, Rendigs, and Welsh, Arthur L. 1981. "The revised test of understanding of college economics," *American Economic Review, Papers and Proceedings, 71* (2), 190-194.

Siegfried, John J., and Fels, Rendigs. 1979. "Research on teaching college economics: A survey," *Journal of Economic Literature, 17* (September), 923-969.

Siegfried John J., and Walstad, William B. 1998. "Research on teaching college economics," in William B. Walstad and Phillip Saunders, eds., *Teaching undergraduate economics: A handbook for instructors* (pp. 141-166). New York: McGraw-Hill.

Welsh, Arthur L., and Fels, Rendigs. 1969. "Performance on the new test of understanding in college economics," *American Economic Review, Papers and Proceedings, 59* (2), 224-229.

APPENDIX 1

Coded List of Schools Providing TUCE-4 Data by Carnegie Classifications

Code Prefix D: Doctoral/Research Universities
Code Prefix M: Master's Colleges and Universities
Code Prefix B: Baccalaureate Colleges
Code Prefix A: Associate's Colleges

School Code	Microeconomics				Macroeconomics			
	Pretest & Posttest	Pretest only	Posttest only	Total	Pretest & Posttest	Pretest only	Posttest only	Total
D401		139		139				
D403		36		36				
D404	68	32	18	118				
D405	150	33	14	197				
D406	170	121	29	320	339	222	71	632
D408	65	25	32	122	263	77	25	365
D409		51		51				
D410	52	54	18	124				
D411	189	37	17	243	56	13	11	80
D413		0	29	29			146	146
D501	64	25	24	113	53	55	10	118
D502					28	18	4	50
D505					17	11		28
D900							98	98
Doctoral	**758**	**553**	**181**	**1,492**	**756**	**396**	**365**	**1,517**
M506						37		37
M601	27	16	7	50	135	118	11	264
M602	18	5		23	28	5	6	39
M603					24	8	5	37
M604	265	149	40	454	152	103	41	296
M605	73	33	8	114				
M606	41	8	1	50				
M607		29		29				
M608	17	5	7	29	32	5	14	51
M609	30	4		34	83	14	13	110
M610	213	48	51	312	128	171	40	339
M611	40	10	3	53				
M612		65		65	21	1	1	23
M613					20	8	2	30
M614	170	15	2	187	103	14	6	123
M615					31	14	1	46
M617	63	8	8	79				
M618	103	7	5	115	126	24	11	161
M619					36	80	55	171
M620	28	8	1	37	17	8	3	28
M622	77	16	12	105	82	16	12	110
M623					136	12	7	155
M624	82	7	1	90	18	5		23

Coded List of Schools Providing TUCE-4 Data by Carnegie Classifications (continued)

School Code	Microeconomics				Macroeconomics			
	Pretest & Posttest	Pretest only	Posttest only	Total	Pretest & Posttest	Pretest only	Posttest only	Total
M625		23		23		50		50
M626	112	67	12	191	24	104	9	137
M628	78	169	4	251		329		329
M629	105	44	165	314				
M630	255	71	34	360	47	13	3	63
M631	26	2	1	29	44	1	7	52
M632					64	15	7	86
M633	27	6	1	34				
M634	23	10	4	37	17	16	1	34
M635	49	12	2	63	20	1		21
M637						74		74
M701	40	11	10	61	20	5	4	29
M702	50	5	2	57				
M703					9	9	4	22
M704					22	2	1	25
M705					45	17	1	63
Masters	**2,012**	**853**	**381**	**3,246**	**1,484**	**1,279**	**265**	**3,028**
B201		24		24	104	67	40	211
B202	37	5		42	82	15	3	100
B203	22	35		57	30	4		34
B204	24	2	1	27	16	2	1	19
B301					16	4	1	21
B302					57	11	5	73
B303	150	52	5	207				
B412					58	10	3	71
Baccalaureate	**233**	**118**	**6**	**357**	**363**	**113**	**53**	**529**
A100	18	5	1	24	32	20	3	55
A103	20	26	2	48	55	88	11	154
A105	24	5	3	32	38	6	2	46
A106	20	5	3	28	61	45	7	113
A107	66	23	2	91				
A108						47		47
A109						28		28
A504	45	14	8	67				
A801	59	19	17	95				
Associate's	**252**	**97**	**36**	**385**	**186**	**234**	**23**	**443**
Total	**3,255**	**1,621**	**604**	**5,480**	**2,789**	**2,022**	**706**	**5,517**

APPENDIX 2

Schools and Instructors Providing TUCE-4 Data

Arkansas State University—Larry Dale
Barry University—Eddie Daghestani, Joan Wiggenhorn
Barton College—John Bethune
Bowling Green State University Firelands—John Girard, Kay Strong
Bridgewater State College—Michael Jones
Butler University—Robert Main, Kathy Paulson Gjerde
Carson-Newman College—Tori Knight, Millicent Sites
Central Arizona College—Bill Demory
Chadron State College—Ronald Burke, Thomas Swanke
Chandler-Gilbert Community College—Douglas Brown
Charleston Southern University—Arnold Hite
College of St. Catherine—Joann Bangs, John Winstandley
CUNY–Hunter College—Purvi Sevak, Mira Tsymuk, Terence Agbeyegbe, Kenneth McLaughlin, Niklas Westelius
Eastern Kentucky University—John Wade
Florida Atlantic University—William Bosshardt, Eric Chiang, Bill Stronge
Florida Southern College—Joan Buccino, Craig Bythewood, Paul Eberle, David Grossman, David Martin
Florida State University—Thomas McCaleb
Fort Hays State University—Kathleen Arano, Emily Breit, Ralph Gamble, Dosse Toulaboe
Frostburg State University—William Anderson, Margaret M. Dalton, Daniel Mizak, John Neral, Anthony G. Stair
Howard University—Emily Blank
Idaho State University—Scott Benson
Indiana University–Kokomo—Fjorentina Angjellari-Dajci, Dmitriy Chulkov
Indiana University Northwest—Donald Coffin
Jamestown Community College—Martha Zenns
John Carroll University—Judith Brenneke, Lindsay Calkins, Lawrence R. Cima, John D. Huffnagle, Simran Kahai, Ida A. Mirzaie, John Soper
Kentucky State University—Albert O. Assibey-Mensah
Kenyon College—Galina An, David Harrington, James Keeler, Jaret Treber
King's College—Margarita Rose
Le Moyne College—Wayne A. Grove
Lyon College—Mahbubul Kabir
Mesa Community College—Douglas Brown, Albert Chavez, Doug Conway, Amy Willis
Mesa State College—Morgan Bridge
Metropolitan State College of Denver—Erick Erickson, Kishore Kulkarni, Leora Starr
Middle Tennessee State University—Maria Edlin
Mississippi State University—Rebecca Campbell, Paul Grimes (field-test coordinator), Meghan Millea, Kevin Rogers
Mount St. Mary's University—John Larrivee, Frank Zarnowski
Mt. Hood Community College—Ted Scheinman
Nassau Community College—Louis Buda, Roberta Schroder
Nichols College—Louise Nordstrom
Oklahoma City University—Becky McMillan, Mary Walker,
Oklahoma State University—Michael Davidsson
Purdue University—Qiangbing Chen, Stefano Gubellini, Bob Holland, Robert Kluin, Daniel Nguyen, Michael Watts (field-test coordinator)
Queens College—Clive Belfield, Hyunbae Chun, Orlando Justo, Kenneth Levin, Farahmand Rezvani, Vamsicharan Vakulabharanam, Thomas J. Webster, Alan Weinman
Quinnipiac University—Mark Gius
Saint Mary's College of California—Kara Boatman, William Lee, Dave Mitchell
Samford University—Jeremy Thornton
San Jose State University—Mark Brady, Doris Cheng, Nicholas Coffaro, Sandra Gullicksen, Edward Lopez, Tom Means, Ninh Nguyen, Ben Parizek, Mike Pogodzinski, Yeung Nan Shieh, Marvin Snowbarger, Edward Stringham, Sean Tanoos
Scottsdale Community College—Douglas Brown
Southeastern Oklahoma State University—Kenneth Chinn
St. Cloud State University—King Banaian, Mary E. Edwards, Nathan Hampton, Richard MacDonald, Ken Rebeck
SUNY–Alfred State College—Fahrettin Dingil
SUNY–Oswego—John Kane
SUNY–Purchase College—Seamus O'Cleireacain
U.S. Naval Academy—Rae Jean Goodman (field-test coordinator), Chris Messineo
University at Buffalo—Peter Morgan
University of Akron—Jay Mutter
University of Alaska Anchorage—Pershing Hill, Stephen Jackstadt, Paul Johnson, S. L. Shapiro
University of Arkansas—David E.R. Gay
University of Colorado, Colorado Springs—John Brock
University of Evansville—Gale Blalock
University of Illinois at Chicago—Helen Roberts, Thomas Smith
University of Memphis—Saktinil Roy
University of Minnesota Duluth—David Doorn
University of Nebraska–Lincoln—Sam Allgood, John Austin, Tammie Fischer, William Walstad
University of North Florida—W. T. Coppedge, Earle Traynham, Louis A. Woods
University of Scranton—Edward Scahill
University of Texas at San Antonio—Ronald Ayers
University of the District of Columbia—Sharron Terrell
Valdosta State University—George Kalcher
Vanderbilt University—Stephen Buckles
Virginia Wesleyan College—David Garraty, Cheul Kang
Weber State University—Doris Geide-Stevenson

APPENDIX 3: Tables A1–A6 for *Unmatched* TUCE-4 Data

TABLE A1. Distribution of Pre- and Posttest Scores on Micro TUCE-4: *Unmatched*

	Pretest (n = 4876)			**Posttest** (n = 3859)		
Raw Score	**No. of Scores**	**Percentile Rank**	**T-Score**	**No. of Scores**	**Percentile Rank**	**T-Score**
30				1		87
29				1		85
28				4		83
27				4		81
26	1		100	11		79
25	1		97	26		77
24	4		94	23	99	74
23	5		91	42	98	72
22	6		88	55	97	70
21	15		85	79	96	68
20	14	99	82	103	94	66
19	19	99	79	108	91	64
18	27	99	76	133	88	62
17	59	98	73	186	85	59
16	73	97	70	203	80	57
15	126	95	67	237	75	55
14	171	93	64	272	69	53
13	251	89	61	311	61	51
12	348	84	58	303	53	49
11	470	77	55	355	46	47
10	542	67	52	338	36	44
9	659	56	49	318	28	42
8	621	43	46	256	19	40
7	549	30	43	187	13	38
6	427	19	40	139	8	36
5	267	10	37	79	4	34
4	131	5	34	55	2	32
3	50	2	31	20	1	30
2	26	1	28	8	0	27
1	14	0	25	2	0	25
Mean Score			9.37			12.59
Std. Deviation			3.35			4.68
Alpha			.47			.70
Std. Error of Measurement			2.45			2.58

TABLE A2. Distribution of Pre- and Posttest Scores on Macro TUCE-4: *Unmatched*

	Pretest (n = 4811)			**Posttest** (n = 3495)		
Raw Score	**No. of Scores**	**Percentile Rank**	**T-Score**	**No. of Scores**	**Percentile Rank**	**T-Score**
30				3		80
29				9		78
28	1		102	18		76
27	1		100	13	99	75
26	0		97	28	99	73
25	1		94	46	98	71
24	4		91	58	97	69
23	8		88	79	95	67
22	9		85	93	93	65
21	12		82	123	90	63
20	18	99	79	131	87	61
19	32	99	77	133	83	59
18	42	98	74	157	79	57
17	67	97	71	178	75	56
16	81	96	68	217	69	54
15	173	94	65	222	63	52
14	216	91	62	241	57	50
13	292	86	59	252	50	48
12	343	80	56	264	43	46
11	466	73	54	266	35	44
10	582	63	51	228	28	42
9	609	51	48	208	21	40
8	547	39	45	200	15	39
7	495	27	42	124	9	37
6	394	17	39	102	6	35
5	238	9	36	58	3	33
4	117	4	33	30	1	31
3	37	1	31	10	0	29
2	22	1	28	1	0	27
1	4	0	25	3	0	25
Mean Score			9.76			14.06
Std. Deviation			3.48			5.28
Alpha			.51			.77
Std. Error of Measurement			2.45			2.53

TABLE A3. Item Analysis: TUCE-4 Micro Pre-Post: *Unmatched*

Item	Correct Answer	Corrected Item—Total Correlation (*n = 3859*)	*Percent Correct* Posttest (*n = 3859*)	*Percent Correct* Pretest (*n = 4876*)
1	A	.26	49%	39%
2	B	.18	39	33
3	A	.23	50	36
4	A	.22	56	14
5	C	.16	45	40
6	C	.30	45	23
7	D	.30	48	44
8	A	.33	36	21
9	D	.18	31	22
10	A	.11	43	36
11	A	.32	31	12
12	C	.24	44	25
13	B	.17	50	37
14	B	.21	44	31
15	C	.18	34	22
16	C	.23	49	42
17	D	.32	42	32
18	B	.21	40	30
19	C	.15	44	43
20	C	.18	30	17
21	D	.24	44	43
22	A	.19	58	54
23	C	.20	31	24
24	B	.21	48	40
25	D	.29	33	24
26	D	.12	34	30
27	B	.16	41	30
28	B	.20	35	24
29	A	.27	37	31
30	D	.30	49	40

TABLE A4. Item Analysis: TUCE-4 Macro Pre-Post: *Unmatched*

Item	Correct Answer	Corrected Item—Total Correlation (*n = 3495*)	*Percent Correct* Posttest (*n = 3495*)	*Percent Correct* Pretest (*n = 4811*)
1	A	.32	53%	24%
2	B	.21	61	50
3	C	.31	68	46
4	D	.33	47	36
5	A	.30	59	12
6	B	.27	45	33
7	B	.26	60	52
8	C	.22	49	40
9	C	.26	34	23
10	B	.19	40	35
11	D	.42	59	34
12	B	.17	55	39
13	B	.31	63	54
14	A	.32	47	25
15	B	.25	60	49
16	C	.16	36	28
17	C	.26	36	31
18	A	.31	45	17
19	A	.22	39	32
20	C	.35	60	51
21	A	.29	42	21
22	D	.35	33	18
23	A	.28	36	27
24	D	.23	32	26
25	B	.36	59	42
26	C	.22	31	23
27	A	.24	32	18
28	D	.31	50	35
29	D	.31	34	24
30	D	.32	43	32

TABLE A5. Percentage Response to Items: TUCE-4 Micro Pre- and Posttests: *Unmatched* (*n = Pre: 4876; Post: 3859*)

Item	A	B	C	D	Blank
1Post	**49***	34	14	3	0.1
1Pre	**39***	45	13	3	0.3
2	6	**39***	23	31	0.6
2	6	**33***	23	37	1.3
3	**50***	10	29	11	0.6
3	**36***	15	36	13	0.9
4	**56***	18	20	6	0.6
4	**15***	30	35	19	1.6
5	32	9	**45***	13	0.8
5	29	11	**40***	19	1.2
6	9	10	**45***	36	0.6
6	13	12	**23***	51	0.8
7	7	17	28	**48***	0.6
7	7	19	29	**44***	1.0
8	**36***	5	3	55	1.1
8	**21***	3	2	73	1.3
9	31	11	26	**31***	0.5
9	32	9	35	**22***	0.7
10	**43***	12	8	37	0.6
10	**37***	10	10	42	1.1
11	**31***	18	22	29	0.8
11	**12***	24	16	47	1.4
12	17	16	**44***	23	0.9
12	13	23	**25***	37	1.8
13	19	**50***	21	10	0.5
13	7	**37***	38	17	1.0
14	19	**45***	15	22	1.0
14	24	**31***	14	30	1.5
15	24	12	**34***	30	0.7
15	27	12	**22***	39	1.4
16	9	30	**49***	11	1.4
16	8	32	**43***	15	1.9
17	13	28	16	**42***	1.7
17	13	26	27	**32***	1.8
18	25	**40***	28	6	0.8
18	29	**30***	32	8	1.1
19	10	14	**44***	31	1.1
19	14	14	**43***	27	1.0
20	11	37	**30***	20	1.7
20	16	49	**17***	17	1.8
21	17	15	22	**44***	2.0
21	14	16	24	**43***	2.5
22	**58***	24	9	7	1.7
22	**54***	20	15	10	2.1
23	18	27	**31***	21	2.6
23	15	35	**24***	23	2.4
24	22	**48***	13	14	2.7
24	26	**40***	16	16	2.2
25	14	23	27	**33***	2.8
25	16	26	31	**24***	2.8
26	13	44	6	**34***	3.2
26	16	42	10	**30***	3.3
27	10	**41***	18	28	3.1
27	16	**30***	21	30	3.0
28	20	**35***	36	6	3.0
28	16	**24***	49	8	3.1
29	**37***	12	39	9	3.2
29	**31***	16	39	11	3.1
30	15	13	19	**49***	4.0
30	18	17	22	**40***	3.9

Note: *Correct answer

TABLE A6. Percentage Response to Items: TUCE-4 Macro Pre- and Posttests: *Unmatched* (*n = Pre: 4811; Post: 3495*)

Item	A	B	C	D	Blank
1Post	**53***	6	9	32	0.6
1Pre	**24***	5	13	57	0.9
2	20	**61***	13	5	0.7
2	26	**50***	16	8	0.8
3	2	13	**68***	17	0.4
3	2	19	**46***	32	0.7
4	11	31	10	**47***	1.1
4	13	43	7	**36***	0.8
5	**59***	6	25	10	0.7
5	**13***	18	44	25	0.9
6	23	**45***	16	15	1.3
6	25	**33***	22	19	1.2
7	15	**60***	17	7	0.6
7	19	**52***	21	8	0.5
8	22	9	**49***	19	0.6
8	21	12	**40***	26	0.5
9	17	29	**34***	19	0.7
9	20	32	**23***	25	0.9
10	28	**40***	21	11	0.6
10	32	**35***	18	15	0.9
11	12	17	11	**59***	1.0
11	18	34	14	**34***	0.6
12	5	**55***	11	28	1.0
12	6	**39***	17	37	0.7
13	14	**63***	13	10	0.7
13	12	**54***	18	15	0.7
14	**47***	16	19	17	0.9
14	**25***	22	29	23	0.9
15	6	**60***	12	21	0.6
15	11	**49***	24	17	0.5
16	21	24	**36***	18	0.9
16	27	27	**28***	17	0.6
17	35	19	**37***	9	0.9
17	38	17	**31***	14	0.4
18	**45***	10	27	18	0.8
18	**17***	20	46	17	0.7
19	**40***	26	19	14	1.1
19	**32***	28	25	14	1.1
20	12	21	**60***	6	1.2
20	18	23	**51***	8	0.7
21	**42***	8	18	30	1.6
21	**21***	9	33	36	1.2
22	24	30	12	**33***	1.8
22	31	35	15	**18***	1.6
23	**36***	26	24	13	1.3
23	**27***	32	24	16	1.4
24	15	28	24	**32***	1.7
24	17	32	24	**26***	1.7
25	12	**59***	15	13	1.3
25	19	**42***	22	15	1.3
26	21	15	**31***	32	1.3
26	27	17	**23***	32	1.5
27	**32***	23	25	20	1.7
27	**18***	22	28	30	1.8
28	18	18	13	**50***	1.7
28	21	23	20	**35***	1.9
29	18	17	29	**34***	1.9
29	19	21	35	**24***	1.9
30	14	23	19	**43***	1.9
30	18	27	22	**32***	2.0

Note: *Correct answer

APPENDIX 4

Microeconomics Test Questions*

Correct options are printed in **boldface**.

1. In an economy where heating oil is the primary source of heat for most households, new supplies of natural gas, a substitute for heating oil, are discovered. Natural gas provides heat at a much lower cost. What is the most likely effect of these discoveries on the market price and quantity of heating oil produced?

	Price	Quantity
A.	**Decrease**	**Decrease**
B.	Decrease	Increase
C.	Increase	Decrease
D.	No change	No change

2. Suppose a city facing a shortage of rental apartments eliminates rent controls. Which of the following is most likely to occur?

   A. a decrease in rents and a decrease in the number of apartment units supplied
   B. an increase in rents and an increase in the number of apartment units supplied
   C. a decrease in the demand for apartments and an increase in the number of apartment units supplied
   D. an increase in the demand for apartments and a decrease in the number of apartment units supplied

3. If all of the firms in a competitive industry are legally required to meet new regulations that increase their costs of production:

   A. supply of the product will decrease.
   B. demand for the product will decrease.
   C. the long-run economic profits of individual firms in the industry will decrease.
   D. the short-run economic profits of individual firms in the industry will increase.

4. At the profit-maximizing level of output, a purely competitive firm will:

   A. produce the quantity of output at which marginal cost equals price.
   B. produce the quantity of output at which marginal cost is minimized.
   C. keep marginal cost lower than price, so profits will be greater than zero.
   D. try to sell all the output it can produce, to spread fixed costs across the largest possible number of units.

5. The demand for a factor of production will usually be more elastic when:

   A. few close substitutes for the factor exist.
   B. the time period under consideration is very short.
   C. demand for the product the factor produces is highly elastic.
   D. the factor's cost is a small part of the final product's total cost of production.

6. Which of the following correctly describes an external benefit resulting from an individual's purchase of flu shots from a doctor?

   A. Doctors earn income by charging for flu shots.
   B. Flu shots are less expensive than catching the flu.
   C. Flu shots reduce the likelihood of others catching the flu.
   D. Flu shots reduce sick days, allowing those who get flu shots to earn more income.

7. A state legislature increased the tax on gasoline sold in the state from $.20 to $.30 per gallon. A supporter said the tax would "make the distribution of after-tax income in the state more equal." This statement would be true only if it could be shown that, after the tax is increased:

   A. people with low incomes buy more gasoline than people with high incomes.
   B. the quantity of gasoline purchased in the state is highly responsive to changes in price.
   C. people with high incomes tend to spend the same proportion of their incomes on gasoline as people with low incomes.
   D. people with high incomes tend to spend a larger proportion of their incomes on gasoline than people with low incomes.

*To order copies of the test, contact the National Council on Economic Education, 1140 Avenue of the Americas, New York, NY 10036, phone (212) 730-7007, or go to www.ncee.net.

8. The opportunity cost of being a full-time student at a university instead of working full-time at a job includes all of the following EXCEPT:

A. payments for meals.
B. payments for tuition.
C. payments for books.
D. income from the full-time job.

9. A recent hurricane destroyed half of the orange crop. Consumers are responding to an increase in the price of oranges by buying more apples. This change is expected to increase the price and quantity of apples sold. In terms of basic supply and demand analysis, there has been a:

A. shift in the demand curve for both oranges and apples.
B. movement along the demand curve for both oranges and apples.
C. shift in the demand curve for oranges and a movement along the demand curve for apples.
D. movement along the demand curve for oranges and a shift in the demand curve for apples.

10. "Water is essential to life, but inexpensive to buy." Which of the following best explains this observation?

A. Water has a high total utility, but a low marginal utility.
B. Water has a low total utility, but a high marginal utility.
C. The quantity supplied of water is less than the quantity demanded at the market price.
D. The quantity supplied of water is greater than the quantity demanded at the market price.

11. The demand for coffee increases and coffee producers begin earning economic profits. Assume the coffee industry is perfectly competitive. Compared to this new situation, in the long run how are the price of coffee and economic profits for coffee producers most likely to change?

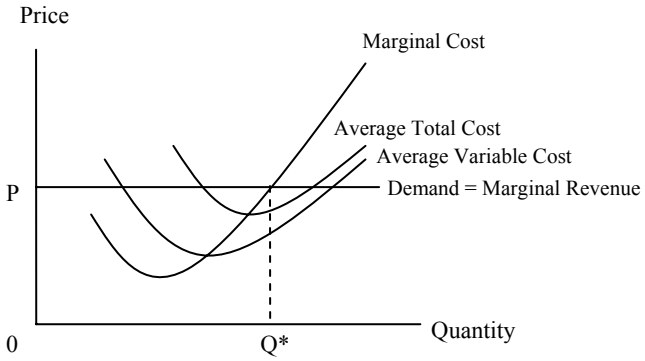
	Price	Economic Profits
A.	**Decrease**	**Decrease**
B.	Decrease	Increase
C.	Increase	Decrease
D.	Increase	Increase

12. A firm is most likely to monopolize a market whenever:

A. it has a U-shaped average total cost curve.
B. fixed capital costs are small relative to total costs.
C. economies of scale are large relative to market demand.
D. the income elasticity of demand is high for the firm's product.

13. As a firm increases its output level in the short run, the costs of producing additional units of output eventually increase because of:

A. diseconomies of scale.
B. diminishing marginal returns.
C. increases in average fixed costs.
D. specialization and division of labor.



14. Which of the following is true for this profit-maximizing firm at price P in the graph above?

A. It is not earning any economic profits.
B. It is currently earning short-run economic profits.
C. It should shut down to minimize its economic losses.
D. It will continue to earn economic profits in the long run.

15. Many U.S. interstate highways are crowded with traffic, but tolls are not collected even when the highways are crowded. Which of the following is true about this no-toll policy?

A. It is efficient because interstates are needed to transport goods.
B. It is efficient because there is no cost of using the interstate once it is built.
C. It is inefficient because each person's use of the interstate adds to the congestion.
D. It is inefficient because tolls would increase government revenues, allowing other taxes to be decreased.

16. "The effect of an excise tax on the products of pollution-producing industries will be a cutback in production. If the tax was levied directly on the amount of pollution generated, the long-run cutbacks in production would be much smaller." This statement is most likely to be:

A. false, provided the amount of the taxes on products and pollution is equal.
B. false, because most firms would rather pay the tax than cut back production.
C. true, because firms would have a greater incentive to adopt new technology that causes less pollution.
D. true, because most taxes levied on pollution affect the demand curve; taxes on products affect the supply curve.

17. A key economic objection to unregulated, profit-maximizing monopoly is that in the short run monopolists:

A. do not try to minimize the costs of the level of output they produce.
B. produce a level of output at which marginal revenue is greater than marginal cost.
C. produce too many products, which they sell at prices that are too high, earning economic profits.
D. restrict output to levels at which their products are valued more than the marginal cost of producing them.

18. A firm is more likely to increase its total revenue by decreasing the price of its product if:

A. the supply of the product is perfectly elastic.
B. there are many close substitutes for its product.
C. the demand for the product is perfectly inelastic.
D. its product accounts for a small portion of a consumer's budget.

19. The market demand for a product has increased if:

A. the product price has increased.
B. more of the product is produced.
C. more of the product can be sold at all possible prices.
D. the cost of producing the product decreased due to new technology.

20. Which of the following statements is correct regarding profit-maximizing firms in the long run?

A. In perfect competition, firms produce an output at which price is less than marginal cost.
B. In perfect competition, firms produce an output at which price is greater than marginal cost.
C. In monopolistic competition, firms produce less than the output at which average total cost is minimized.
D. In monopolistic competition, firms produce more than the output at which average total cost is minimized.

21. One way in which monopolistic competition and oligopoly are similar is that, typically, in both kinds of industries:

A. there are no barriers to entry for firms.
B. each firm has a small share of the market for a product.
C. there are a large number of independent firms selling similar but differentiated products.
D. resources are underallocated to the production of goods and services produced by these firms at their profit-maximizing level of output.

22. Suppose the only two cola companies (Acola and Bcola) in a nation are deciding whether to charge high or low prices for their colas. The companies' price strategies are shown in the table below. The four pairs of payoff values show what each company expects to earn or lose in millions of dollars, depending on what the other company does.

		Acola's Price Strategy	
		High Price	**Low Price**
Bcola's Price Strategy	**High Price**	Acola +$100 Bcola +$100	Acola +$250 Bcola -$50
	Low Price	Acola -$50 Bcola +$250	Acola +$50 Bcola +$50

If both companies believe that most consumers are soon going to quit drinking colas, and switch to other products, what is the equilibrium outcome?

A. Both Acola and Bcola will charge a low price.
B. Both Acola and Bcola will charge a high price.
C. Acola will charge a low price; Bcola will charge a high price.
D. Acola will charge a high price; Bcola will charge a low price.

23. "Ticket prices for professional team sports are high because the team owners just pass the costs of the athletes' high salaries on to ticket buyers." Is this statement generally correct or incorrect? Why?

A. Correct. High sports salaries contain "economic rent" and economic rent normally gets passed on to consumers.
B. Correct. High sports salaries force owners to charge high ticket prices, which they can pass on to consumers because demand is elastic.
C. Incorrect. High sports salaries contain "economic rent" and would not be so high if the public were unwilling to buy tickets at the high prices.
D. Incorrect. Owners can afford to pay the high salaries without raising ticket prices. They raise prices simply to increase their marginal revenue above their marginal cost.

24. In Sunshine City, one local ice cream company operates in a competitive labor market and product market. It can hire workers for $45 a day and sell ice cream cones for $1.00 each. The table below shows the relationship between the number of workers hired and the number of ice cream cones produced and sold.

Number of Workers Hired	Number of Ice Cream Cones Sold
4	340
5	400
6	450
7	490
8	520

As long as the company stays in business, how many workers will it hire to maximize profits or minimize losses?

A. 5
B. 6
C. 7
D. 8

25. Government decisions that are more likely to suffer from the influence of special interest groups are typically ones that yield:

A. costs to all now and benefits to all later.
B. benefits to all now and costs to all later.
C. large benefits for each individual in a large group and small losses for each individual in a small group.
D. large benefits for each individual in a small group and small losses for each individual in a large group.

26. In a country where only two goods are produced and consumed, the production and consumption of Good X results in external benefits, while the production and consumption of Good Y results in external costs. Would unregulated markets produce too much or too little of Good X and Good Y, compared to the efficient output levels for these products?

	Good X	Good Y
A.	Too much	Too much
B.	Too much	Too little
C.	Too little	Too little
D.	**Too little**	**Too much**

27. Public goods are generally provided by government rather than private firms because:

A. people must pay for public goods if they want to consume them.
B. public goods can be used by one person without reducing the amount that is available to others.
C. special interest groups get the government to produce public goods, even if the costs of producing them are greater than the benefits.
D. it is less expensive for government to produce goods that are most important to consumers because the government does not make profits.

28. The table below shows the tons of rice and corn that can be produced in Country X and Country Y in one year, using the same amount of productive resources.

	Rice	Corn
Country X	20	10
Country Y	16	4

According to the theory of comparative advantage, what should firms in Country X do?

A. export rice to Country Y and import corn
B. export corn to Country Y and import rice
C. export both rice and corn to Country Y
D. import both rice and corn from Country Y

29. "To correct our balance of trade deficit, we should increase tariffs on imported goods." If tariffs are increased, the long-run effect is most likely to be:

A. a decrease in both U.S. imports and exports.
B. an increase in both U.S. imports and exports.
C. a decrease in U.S. imports, and an increase in U.S. exports.
D. an increase in U.S. imports, and a decrease in U.S. exports.

30. If the exchange rate between dollars ($) and yen (¥) changes from $1 = ¥200 to $1 = ¥100, and domestic prices in both countries stay the same, has the dollar appreciated or depreciated, and would U.S. imports from Japan become less expensive or more expensive?

Value of the dollar	U.S. imports from Japan
A. Appreciated	Less expensive
B. Appreciated	More expensive
C. Depreciated	Less expensive
D. Depreciated	**More expensive**

APPENDIX 5

Macroeconomics Test Questions*

Correct options are printed in **boldface**.

1. Which of the following is classified as investment in national income (GDP) accounting?

   A. building a new factory
   B. buying a 10-year-old house
   C. depositing money in a bank
   D. purchasing corporate stocks and bonds

2. The consumer price index in an economy is 180 one year and 189 the next year. The rate of inflation in the economy over that year period is:

   A. 1 percent.
   B. 5 percent.
   C. 8 percent.
   D. 18 percent.

3. In the short run, how will an increase in aggregate demand most likely affect the overall price level and real GDP?

Price Level	Real GDP
A. Decrease	Decrease
B. Decrease	Increase
C. Increase	**Increase**
D. Increase	Decrease

4. The limit of total productive capacity in an economy is set by:

   A. the amount of money in circulation.
   B. business demand for goods and services.
   C. the amount of government spending and taxation.
   D. the quantity and quality of its productive resources.

5. The basic money supply (M1) in the United States consists primarily of:

   A. currency and checkable deposits.
   B. currency and government bonds.
   C. currency, checkable deposits, and government bonds.
   D. currency, checkable deposits, and credit card accounts.

6. Which of the following actions by the Federal Reserve would have similar effects on the size of the U.S. money supply?

   A. decreasing the reserve ratio and selling government securities
   B. increasing the reserve ratio and selling government securities
   C. decreasing the discount rate and selling government securities
   D. increasing the discount rate and buying government securities

7. A monetary policy will increase GDP in the short run if:

   A. interest rates increase, encouraging more saving.
   B. interest rates decrease, encouraging more investment.
   C. personal savings increase to finance present consumption.
   D. personal savings decrease to finance future consumption.

8. Which of the following best explains the statement "every government has a fiscal policy, whether it realizes it or not"?

   A. Every government is forced to do something about recessions and inflation, whether it wants to or not.
   B. In many cases, decisions to spend money must be made even though the expenditure runs contrary to the policy indicated.
   C. Every government sets tax and expenditure programs, which influence economic activity and the components of GDP.
   D. Every government makes decisions about the quantity of money in the economy, which influence credit conditions and the rate of interest.

9. Increased government budget deficits cause crowding out if:

   A. imports are decreased more than exports.
   B. a recession causes businesses to lower prices or shut down.
   C. private investment spending for capital goods is decreased.
   D. spending on projects funded by the deficit increases households' spending on goods and services.

**To order copies of the test, contact the National Council on Economic Education, 1140 Avenue of the Americas, New York, NY 10036, phone (212) 730-7007, or go to www.ncee.net.

10. Reducing inflation will require the monetary authorities to make larger changes in the money supply if most people expect a rapidly rising price level because:

A. inflation reduces the opportunity cost of holding money.
B. inflation increases the opportunity cost of holding money.
C. there is a tradeoff between unemployment and inflation in the short run but not in the long run.
D. there is a tradeoff between unemployment and inflation in the long run but not in the short run.

11. Use this information to answer the next question. All numbers are in millions.

Population of the nation 125
Size of the labor force 75
Number of employed workers..... 50
Number of unemployed workers..... 25

What is the unemployment rate for this nation?

A. 5 percent
B. 20 percent
C. 25 percent
D. 33 percent

12. If the price level is expected to increase by three percent next year and a key market interest rate is seven percent, the real rate of interest is:

A. three percent.
B. four percent.
C. seven percent.
D. ten percent.

13. Suppose an economy in long-run equilibrium experiences a supply shock from substantially higher energy costs. In which of the following ways are real GDP and the price level most likely to change?

	Real GDP	Price Level
A.	Decrease	Decrease
B.	**Decrease**	**Increase**
C.	Increase	Decrease
D.	Increase	Increase

14. Which of the following is true in the short run when comparing an increase in government spending to an increase in private investment spending?

A. They will both increase aggregate demand.
B. Government spending is inflationary; private investment spending is not.
C. Government spending must equal taxes; private investment spending must equal saving.
D. The increase in investment spending will result in a greater increase in employment than the increase in government spending.

15. In the short run, aggregate demand in a country will increase if there is an increase in the:

A. tax rates in the country.
B. money supply of the country.
C. prices of resources in the country.
D. level of technology in the country.

16. How will market interest rates and bond prices most likely change if the Federal Reserve decides to make a small, one-time increase in the money supply?

	Market Interest Rates	Bond Prices
A.	Increase	Increase
B.	Increase	Decrease
C.	**Decrease**	**Increase**
D.	Decrease	Decrease

17. Which of the following would most likely result if the federal government increased spending without increasing tax revenues during a period of full employment?

A. a recession
B. a decrease in interest rates
C. an increase in the price level
D. a decrease in the national debt

18. Which of the following actions by a nation's central bank would be most effective in reducing inflation?

A. selling government securities on the open market
B. lowering margin requirements on purchases of financial securities
C. reducing the rate of interest it charges on loans to commercial banks
D. reducing reserve requirements on deposits held by commercial banks

19. Over a five-year period GDP in a nation increased from $10 trillion to $15 trillion, while the GDP price deflator increased from 100 to 125. Approximately how much is GDP in year five, stated in terms of year-one dollars?

A. $12 trillion
B. $14 trillion
C. $16 trillion
D. $19 trillion

20. Which of the following best explains why a $7 billion tax cut can lead to a $9 billion increase in consumer spending in the short run?

A. Tax cuts reduce government spending, which encourages consumer spending.
B. Tax cuts reduce interest rates, which stimulates consumer spending and borrowing.
C. Tax cuts increase disposable income, which leads to higher national income and additional consumer spending.
D. Tax cuts increase government transfer payments, which leads to higher national income and additional consumer spending.

21. Which of the following is most important in increasing a nation's economic growth in the long run?

A. higher rates of technological change
B. higher levels of government spending
C. increasing exports and decreasing imports
D. increasing consumer spending in the economy

22. Suppose commercial banks have no excess reserves. Then new deposits totaling $1 billion come into the banking system. The required reserve ratio is 20 percent. What is the maximum amount by which banks can increase deposits in the entire banking system?

A. $0.5 billion
B. $2.0 billion
C. $2.5 billion
D. $5.0 billion

23. In the short run, compared to an increase in government spending, a decrease in federal taxes will cause aggregate spending in the economy to change in the:

A. same direction, but encourage private over public spending.
B. opposite direction, and encourage private over public spending.
C. same direction, but have neutral effects on private and public spending.
D. opposite direction, and have neutral effects on private and public spending.

24. Assume that the economy is at full employment and is experiencing rapid inflation. Which of the following combinations of monetary and fiscal policies would reduce inflation most, assuming the dollar values for both policy changes are the same amount?

	Monetary Policy	Fiscal Policy
A.	Buy government securities	Increase the federal budget deficit
B.	Buy government securities	Decrease the federal budget deficit
C.	Sell government securities	Increase the federal budget deficit
D.	**Sell government securities**	**Decrease the federal budget deficit**

25. Actual GDP in a country is estimated to be 10 percent below potential GDP. Prices are virtually unchanged from one year ago. Unemployment is 12 percent of the civilian work force, much higher than it has been in many years. Which of the following policies would be the most appropriate for improving these economic conditions?

    A. reductions in the federal debt
    B. decreases in interest rates by the central bank
    C. increases in corporate and personal income taxes
    D. increases in reserve requirements on deposits at commercial banks to protect depositors

26. If workers or businesses anticipate that an expansionary monetary policy will increase inflation, the effects of this policy on real output will be:

    A. smaller, if businesses lower prices on the products they make.
    B. larger, if businesses lower prices on the products they make.
    C. smaller, if workers demand and receive higher wages.
    D. larger, if workers demand and receive higher wages.

27. "For the past fifteen months, unemployment has been under 5 percent. Consumer prices increased by 2 percent over the level a year ago. Total production of goods and services is projected to be 5 percent higher this year than it was last year."

    Which of the following policies would be most appropriate for short-run stabilization objectives?

    A. relying on automatic economic stabilizers
    B. increasing both personal and corporate income taxes
    C. passing new corporate tax incentives to encourage investment
    D. increasing the minimum wage and expanding the number of jobs covered by minimum wage laws

28. Which of the following is most likely to occur when firms in other countries want to build factories in the United States or purchase U.S. financial securities?

    A. The demand for foreign currencies in the U.S. will decrease and the dollar will appreciate.
    B. The demand for foreign currencies in the U.S. will increase and the dollar will depreciate.
    C. The demand for the dollar will decrease in other countries and the dollar will depreciate.
    D. The demand for the dollar will increase in other countries and the dollar will appreciate.

29. A small country that has experienced high inflation for the past decade decides to set the value of its currency equal to the value of a currency in a large nation that has had very low inflation for the past 50 years. The small country benefits because this action:

    A. gives its central bank more flexibility in using monetary policies to reduce the rate of inflation.
    B. makes it easier for the national government to use monetary policy and fiscal policy to fight unemployment.
    C. promotes higher levels of international trade by reducing taxes on imports and increasing subsidies for exports.
    D. establishes greater confidence among domestic and international investors that the country's inflation will be brought under control.

30. If income and consumption in the U.S. economy are growing faster than in the economies of the nations that are its major trading partners, U.S. imports are most likely to:

    A. decrease less than U.S. exports.
    B. decrease more than U.S. exports.
    C. increase less than U.S. exports.
    D. increase more than U.S. exports.

APPENDIX 6

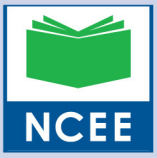
Sample Answer Sheet Marked with Scoring Key for Microeconomics Questions

	A	B	C	D		A	B	C	D		A	B	C	D
1	●	○	○	○	**11**	●	○	○	○	**21**	○	○	○	●
2	○	●	○	○	**12**	○	○	●	○	**22**	●	○	○	○
3	●	○	○	○	**13**	○	●	○	○	**23**	○	○	●	○
4	●	○	○	○	**14**	○	●	○	○	**24**	○	●	○	○
5	○	○	●	○	**15**	○	○	●	○	**25**	○	○	○	●
6	○	○	●	○	**16**	○	○	●	○	**26**	○	○	○	●
7	○	○	○	●	**17**	○	○	○	●	**27**	○	●	○	○
8	●	○	○	○	**18**	○	●	○	○	**28**	○	●	○	○
9	○	○	○	●	**19**	○	○	●	○	**29**	●	○	○	○
10	●	○	○	○	**20**	○	○	●	○	**30**	○	○	○	●

APPENDIX 7

Sample Answer Sheet Marked with Scoring Key for Macroeconomics Questions

	A	B	C	D
1	●	○	○	○
2	○	●	○	○
3	○	○	●	○
4	○	○	○	●
5	●	○	○	○
6	○	●	○	○
7	○	●	○	○
8	○	○	●	○
9	○	○	●	○
10	○	●	○	○
11	○	○	○	●
12	○	●	○	○
13	○	●	○	○
14	●	○	○	○
15	○	●	○	○
16	○	○	●	○
17	○	○	●	○
18	●	○	○	○
19	●	○	○	○
20	○	○	●	○
21	●	○	○	○
22	○	○	○	●
23	●	○	○	○
24	○	○	○	●
25	○	●	○	○
26	○	○	●	○
27	●	○	○	○
28	○	○	○	●
29	○	○	○	●
30	○	○	○	●



National Council on Economic Education

1140 Avenue of the Americas, New York, NY 10036
212.730.7007 *fax* 212.730.1793 econed@ncee.net www.ncee.net

1-56183-609-5