Econ-assessments.org: Automated Assessment of Economics Skills

Douglas McKee¹ · Steven Zhu¹ · George Orlov¹

Published online: 27 October 2022
© EEA 2022

Abstract
Standard assessments can give instructors reliable objective measures of student skills at the beginning and end of a term. We describe seven standard assessments that we have developed for commonly taught economics courses, and introduce a new website, econ-assessments.org, that allows instructors worldwide to set up any of our assessments for their students. Instructors get a link to their chosen assessment that can be shared with students, and students take the timed assessment when they follow the link. At the end of the test-taking period, instructors receive a report that summarizes performance in aggregate and by learning goal.

Keywords Economics · Assessment · Concept inventory

JEL Classification A2 · C1

Introduction
High quality standard assessments of student knowledge and skills can be incredibly valuable to instructors, researchers, and students. A standard assessment allows instructors to measure student learning in their course across semesters to estimate the impact of any teaching changes they may have implemented. Instructors can look at average achievement, gaps between demographic groups, or performance on individual learning goals. They can also use assessments at the beginning of the term to see how well-prepared their students are for the course and to identify students that might benefit from extra attention.

¹ Department of Economics, Cornell University, Ithaca, USA

Douglas McKee
douglas.mckee@cornell.edu

Steven Zhu
sz297@cornell.edu

George Orlov
george.orlov@cornell.edu
Standard assessments have become popular in other STEM disciplines as a tool for researchers to measure the impact of pedagogical innovations on student learning. At the time of writing, the Discipline-Based Science Education Research Center (DB-SERC) at the University of Pittsburgh lists 58 such assessments for Physics, 33 for Biology, 28 for Chemistry, 14 for Computer Science, and 5 for Mathematics. These assessments have enabled great progress in discipline-based education research. For example, researchers have used standard assessments to provide classroom-based evidence for the effectiveness of active learning (Freeman et al. 2014), peer instruction (Crouch and Mazur 2001), and the determinants of race and gender-based performance gaps (Salehi et al. 2019; Bottan et al. 2022). This research in turn greatly benefits students, as instructors make use of new evidence-based practices to improve student performance, reduce achievement gaps, and potentially increase diversity of the undergraduate economics student population.

Until very recently, the only standard assessment appropriate for undergraduate economics courses was the Test of Understanding College Economics (TUCE) which tests introductory economics knowledge (Walstad and Rebeck 2008). The first version of the TUCE was published in 1967, but it has been revised several times since then, and the latest revision was released in 2007. The TUCE includes questions based on content commonly taught in introductory microeconomics and macroeconomics courses. Economics instructors have used the TUCE informally to measure learning in their classrooms and more formally to publish research on a variety of teaching innovations (e.g., Becker and Powers 2001; Coates et al. 2004; Emerson and Taylor 2004). We believe our new assessments and our automated system can enable even more instructors to carefully evaluate their teaching and answer important education research questions.

In the next section, we provide an overview of the Cornell Suite of Economics Assessments, which currently includes tests appropriate for several commonly taught undergraduate courses. We then describe econ-assessments.org, a website with a fully automated system that allows instructors to give online versions of assessments to their students and generates detailed performance reports. We give advice on how instructors and researchers can best integrate these assessments into their courses in the following section, and we conclude by describing our future plans to extend econ-assessments.org with additional assessments and functionality.

The Cornell Suite of Economics Assessments

The Cornell Suite of Economics Assessments is currently comprised of seven tests that evaluate the knowledge and skills of economics students. Five of the assessments evaluate skills taught in undergraduate economics courses, while two measure the math skills required by these courses. The test questions are multiple-choice to enable reliable unbiased automated grading, and they measure both understanding of core concepts and several higher order skills. Much research has shown that carefully designed multiple-choice questions can assess higher order skills as effectively as short-answer or essay questions (Wainer and Thissen 1993; Walstad 1998;
Students are given between 35 and 60 min to complete an assessment. Development of the Cornell Suite of Economics Assessments started as part of a project to incorporate active learning methods into the undergraduate curriculum at Cornell University (Cornell Center for Teaching Innovation 2022). The primary goal of that project was to increase student learning, and in order to evaluate the project’s success, instructors needed high quality measures of student learning in each course. Project personnel developed each assessment using a process based on the procedure described in Adams and Wieman (2011). Developers started by documenting the learning goals they wanted to test in the assessment and drafting a set of corresponding multiple-choice questions. Faculty were recruited inside and outside Cornell to provide feedback on whether the assessment effectively evaluated expert-level thinking. Following revisions, interviews were conducted with undergraduate students who had previously taken the relevant course where they verbalized their thought process as they answered the questions. Based on these interviews, questions were clarified and potential answers that corresponded to common mistakes were added. Our development and validation processes are described in detail in McKee and Orlov (2022).

All seven of the assessments now available on econ-assessments.org are in active use at Cornell, and several have already been piloted in classrooms at other institutions. We have received positive feedback on the utility of these assessments from nearly all instructors who have participated. Some instructors have used the assessments to identify parts of their courses which they can improve, and others have used them to evaluate the impact of teaching changes year-over-year. A prominent example of the latter was an analysis of the impact of the COVID-19 pandemic on student learning. In Orlov et al. (2021a), we collaborate with three Cornell instructors and three economics instructors from other institutions and compare student scores on four of our assessments before and after the pandemic to investigate which students were most affected by the switch to remote instruction and how different teaching methods mitigated these effects.

Mathematics for Economics Skills Assessment-Foundations (MESA-Foundations)

MESA-Foundations is a 31-question multiple-choice test designed to assess the mathematical skills that are required for success in introductory economics courses. It contains questions that test skills in arithmetic, geometry and graphs, and algebra. The test helps instructors teach students in their economics courses based on the skills they have instead of the skills we hope they have. In addition, the test can identify early in the term those students who will need extra support during the term, and it has been shown to be a strong predictor of performance in introductory courses (Orlov et al. 2021b). MESA-Foundations was co-developed by Irene Foster (George Washington University), George Orlov (Cornell), and Doug McKee (Cornell), and has been piloted at Cornell University and George Washington University.
Mathematics for Economics Skills Assessment-Intermediate (MESA-Intermediate)

MESA-Intermediate is a 37-question multiple-choice test designed to assess the mathematical skills that are required for success in intermediate-level economics courses. It contains questions that test skills in geometry and graphs, algebra, and calculus. Like MESA-Foundations, this test helps instructors teach students based on their true skill levels. MESA-Intermediate was co-developed by Irene Foster (George Washington University), George Orlov (Cornell), and Doug McKee (Cornell), and has been piloted at Cornell, George Washington, and the Copenhagen Business School.

Principles of Economics Skills Assessment-Microeconomics (PESA-Micro)

PESA-Micro is designed to test the learning goals of a typical introductory course in microeconomics. It is composed of 30 multiple-choice questions. We have given this assessment in several Cornell classes as both an end-of-semester assessment for our introductory course and a test of pre-requisite skills in our intermediate-level courses. PESA-Micro is an alternative to the Test of Understanding College Economics (TUCE) that is appropriate for courses that are more explicitly theory-based.

Intermediate Economics Skills Assessment-Microeconomics (IESA-Micro)

IESA-Micro is designed to test the learning goals of a calculus-based course in intermediate microeconomic theory. It is composed of 31 multiple-choice questions that test both conceptual understanding and the ability to use economic theory to make predictions. IESA-Micro has been piloted at Cornell.

Economic Statistics Skills Assessment (ESSA)

ESSA is designed to test the learning goals of an introductory probability and statistics course that is geared toward economics students. It is composed of 20 multiple-choice questions, and students may refer to a provided formula sheet during the test. ESSA has been piloted extensively inside and outside Cornell. McKee and Orlov (2022) documents in detail the design, development, and validation of this assessment. ESSA is also the first assessment on the site to be available in Spanish.

Applied Econometrics Skills Assessment (AESA)

AESA is designed to test the learning goals of an applied econometrics course that follows an introductory course in probability and statistics. It is composed of 24 multiple-choice questions about linear regression models and more advanced
methods including binary outcome models, instrumental variables, and fixed effects. AESA has been piloted inside and outside Cornell.

**Applied Econometrics Skills Assessment-Core Skills (AESA-Core)**

AESA-Core is designed to assess students’ facility with linear regression models. It is composed of the same 16 multiple choice questions about linear regression models that are included in the broader Applied Econometrics Skills Assessment (AESA). The AESA-Core questions have been piloted inside and outside Cornell.

**The Automated Assessment System**

Econ-assessments.org provides substantial information about the Cornell Suite of Economics Assessments and the services offered. It explains the benefits of giving economics students standard assessments and describes each of the seven assessments that are currently available. Instructors can immediately download the learning goals that are evaluated by each assessment to compare them with the learning goals of their own courses to determine which assessments might be useful. Contact information for instructors who have experience using these assessments with their students is also available upon request.

Instructors can request questionnaires by filling out a short form and providing their contact information, including their institutional email address and web page. Because we do not want students to be able to access the questions before they take the assessment, we validate instructors before distributing questionnaires. After we manually verify the instructor’s identity, the system stores the validation information for future use and sends any questionnaires that were requested to the instructor.

Once the instructor decides they want to set up an assessment for their students, they can fill out another short form where they provide their contact information and select a specific assessment. The instructor then provides some basic information about the course and chooses a date and time when the assessment will close. If the instructor requires students to take the assessment or is planning to give extra credit to students who complete it, they can ask to receive a list of participating students when the assessment is over.

Instructors have quite a bit of flexibility in how they field their assessment. They can set up one assessment and share the link with students in multiple classes as long as those classes teach similar content and the timing of the assessment is the same. Alternatively, instructors can set up separate assessments for each of several sections. In the latter case, the instructor will receive separate performance reports and student participation lists for each section.

Based on the information entered by the instructor, econ-assessments.org creates a Qualtrics survey that includes the appropriate assessment questions and sends a link for this survey to the instructor. The instructor can distribute this link to their class when they want to open the survey. The survey begins by displaying the following Cornell IRB-approved consent statement:
You are invited to participate in a collaborative research project led by Dr. Douglas McKee, a faculty member in the Department of Economics at Cornell University. The purpose of this research is to measure student understanding of [XXX] methods used by economists.

**What we will ask you to do:** This survey poses a number of questions about concepts and methods used by economists. The primary (but not only) purpose of the survey is for your instructor(s) and department to evaluate the effectiveness of their courses. This consent form asks for your permission to use your responses for additional research purposes.

**Risks and discomforts:** We do not anticipate any risks from participating in this research.

**Benefits:** The benefits of participating in this research are the opportunities to learn, practice some new material, and contribute to future instructional innovations in economic education.

**Compensation for participation:** You will either receive course credit or no compensation for your participation. If your instructor has assigned course credit for completing this survey, you must provide your name and school email address at the end of the survey to receive this credit.

**Privacy/Confidentiality/Data Security:** Your identifying information will be kept with your survey data for the purposes of matching survey responses. Your instructor will have access to your responses at the end of the semester/quarter. Your identity will never be associated with your responses publicly.

**Taking part is voluntary:** You may choose whether to consent to having your survey data used for research purposes. Participation credit (if your instructor has assigned such credit) does not depend on any specific answers you provide in the survey or on whether you consent to having your answers used in research. You may complete this survey for your instructor in order to obtain participation credit in your course (if applicable).

**If you have questions:** If you have any questions about this study, please contact Dr. Douglas McKee (douglas.mckee@cornell.edu) or Dr. George Orlov (george.orlov@cornell.edu).

**Statement of consent:**
- I have read and understood the above information. I consent to take part in the study.
- I do not consent to take part in the study.

**Reminder:** If your instructor has assigned course credit for completing this survey, you must provide your name and school email address at the end of the survey to receive this credit, independent of consenting to participate in the research study.

If you have any questions or concerns regarding your rights as a subject in this study, you may contact the Cornell Institutional Review Board (IRB) for Human Participants at 607-255-6182 or access their website at [http://www.irb.cornell.edu](http://www.irb.cornell.edu). You may also report your concerns or complaints anonymously through Ethicspoint online at [www.hotline.cornell.edu](http://www.hotline.cornell.edu) or by calling toll free...
at 1-866-293-3077. Ethicspoint is an independent organization that serves as a liaison between the University and the person bringing the complaint so that anonymity can be ensured.

All students, regardless of how they answer the consent question, are then given the assessment. Data from students who choose not to take part in the study are removed before performance reports are generated or any samples are analysed for research purposes. All students are included in the participant lists sent to the instructor.

Twenty-four hours before an assessment window closes, the system emails the instructor with the current number of responses and suggests that the instructor send a reminder to students. Instructors are also given the option to change the end date of the assessment before the assessment closes.

When the assessment ends, econ-assessments.org closes the Qualtrics survey and generates a performance report for the instructor. This report includes the response rate of students in the class, a histogram showing the overall score distribution, and several summary performance statistics. The mean, standard deviation, and quartiles are included for the overall score as well as subscores based on questions grouped by learning goals. The report includes a copy of the assessment’s learning goals for reference. If the instructor is giving their students course credit for completing the assessment, the system also emails them a .csv file containing the names and campus email addresses of participating students.

If an instructor requires individual-level performance and demographic data for a research project and has local IRB approval, we will send them the raw Qualtrics student response data. This includes the time each student spent on the assessment and allows instructors to discard observations of students that spent an unreasonably short time on the test. We currently include all consenting students in the performance reports but are seriously considering removal of these low-effort students in the future. We have found in our own research that their removal improves reliability of the data and yields a more normal distribution of scores.

**Using Econ-assessments.org**

The details of how instructors give these assessments to their students can make a big difference in student participation rates, the effort students put in, and the quality of the data collected.

We recommend letting students know early in the term how and why you will be incorporating the assessment into the course. It’s good to include this information in the syllabus of the course as well as talking about it during class. If you are testing their skills at the beginning of the term, tell students that this will allow you to teach them based on what they actually know instead of what they “should” know. You can use the results to review appropriate material and fill missing knowledge gaps before it’s too late. Similarly, a beginning-of-term test tells students what skills they need to improve in order to succeed in the class.
Standard assessments given at the end of the term before final exams provide similar benefits, telling instructors what to review for the class and telling students what they need to study. We often tell our students that their effort has a payoff for future students, as the data we collect will help us evaluate the effectiveness of any changes we made to our teaching in the current term.

Students can also be informed that putting a good effort into a standard assessment can have a big positive effect on their performance on exams in the class. We encourage instructors to share evidence of the “practice effect” with students such as Rawson and Dunlosky (2012), who review recent research and report that “the modal finding is that performance on a subsequent final test is greater following a practice test versus no practice test” (p. 420).

The best way we have found to encourage students to take the assessment is to give them course credit for participation. This can take the form of simply requiring participation, making it a small portion of their grade, or giving “extra credit.” Our experience is that just one or two extra points on an exam score can make a huge difference.

Another way to encourage serious participation is to explain that students played a large part in the creation and validation of the assessment that they will be taking. Specifically, students took early draft versions of the assessment, verbalizing their thought process while answering the questions. This allowed the developers to clarify many of the questions and include more plausible incorrect answer choices. When students understand that the questions in the assessment have been evaluated by their peers, they are more likely to appreciate the validity of the assessment process.

Instructors that use econ-assessments.org are not allowed to consider students’ scores when giving credit to students. We impose this requirement to eliminate (or at least reduce) the incentive of students to find solutions on the internet or cheat in other ways. This behaviour would obviously invalidate the assessment as a measure of student knowledge and skills.

The faculty we have worked with have found that our online assessment system requires very little instructor work beyond filling out the initial request form, sending the link to their students, and reminding their students when the test-taking window is about to close. When we originally fielded our assessments as classroom tests with bubble sheets, our instructors were quite unhappy giving up precious class time for the test-taking. The time it took to process the bubble sheets and manually generate reports was also not appreciated. The econ-assessments.org system requires zero class time, and faculty automatically receive reports within 15 min of the close of the test-taking window. While some instructors believe that student effort is higher for an in-person test, our experience is that effort levels seem to differ very little between in-class and online assessment. In addition, the online assessment has the advantage that you can recognize some of the students putting in low effort by observing the time they spent on the assessment.
Future Work

While we believe econ-assessments.org in its current form provides a valuable service to the economics teaching and research community, we are planning to make several improvements over the next year to the assessments and the site itself.

Perhaps the most important improvement is the implementation of a process for incorporating instructor feedback on assessment questions and the learning goals that are evaluated. Economics is an evolving discipline, and what we teach at the undergraduate level changes over time to reflect this. To remain relevant, our assessments must also change to evaluate these new learning goals. We view the development of our assessments as an ongoing process of refinement and improvement driven by the teaching community. We are planning to form a review board composed of instructors around the world for each assessment that will consider suggestions from the community and analyse the data we collect on the site to make sure each assessment is as accurate and reliable as possible. Instructors will be able to choose to use the latest version of an assessment or an older version if they want perfect comparability to data they collected earlier.

Second, we are planning to incorporate additional assessments into the system. At Cornell, we are piloting assessments of introductory macroeconomics (PESA-Macro) and intermediate macroeconomics (IESA-Macro) that we are planning to share with the community before the end of 2022. Because there is less consensus about what is taught in macroeconomics courses relative to microeconomics or econometrics, feedback from the community will be crucial as these assessments evolve and become more useful to a wide range of macroeconomics instructors. We also have math-intensive versions of our introductory economic statistics assessment (MI-ESSA) and a theory-based econometrics skills assessment (TESA) that will be made available to outside instructors in the near future. We are also very interested in working with other instructors on new economics assessments or incorporating existing high-quality assessments developed outside Cornell.

The third major change we have planned is to extend the current reporting system. In addition to the static PDF reports that we deliver now, we will provide interactive reports that let instructors compare current results for a course to historical data. Instructors will know where their students stand in comparison to other students around the world who have taken the assessment. We will construct these benchmark averages when we have sufficient data from a wide range of economics courses. Perhaps more important, we will make it easier for instructors to compare current performance of students to those who have taken the course with them in the past. Specifically, instructors will be able to see where students are learning more (or less) than the comparison group for each topic. This information will help instructors identify areas for improvement in their course and the impact of changes they make to their teaching. The new system will also allow instructors to see performance of
different subgroups of students (for example, by race, gender, major, class year) so they can identify which groups might benefit from extra attention.¹

Conclusion

Education research communities in multiple STEM disciplines have greatly benefited from having access to rigorously developed standard assessments of course-specific learning goals, and until now, no assessments of this kind have existed for economics courses beyond the introductory level. In this paper, we describe the Cornell Suite of Economics Assessments that lets instructors and researchers evaluate a wide range of skills that they either require in their courses or teach during the term. We also describe a new website, econ-assessments.org, that automates the process of giving students these assessments and generates detailed reports of student performance.

We hope that econ-assessments.org will inspire instructors to do more formal evaluation of changes in pedagogy that can be shared with the greater economics and STEM community through publication. Standard easy-to-grade assessments of explicitly defined learning goals have enabled a steady stream of research on teaching in other STEM disciplines, and we believe the same can and should happen in economics.

Acknowledgements The software behind econ-assessments.org is based in large part on work done by Natasha Holmes and Cole Walsh for the PLIC and BIO-MAPS assessment sites. Their work in turn builds on that of the Lewandowski Group at the University of Colorado to support online administration of the E-CLASS assessment. We would also like to thank the many faculty inside and outside Cornell who have provided feedback throughout the development and validation of these assessments as well as the instructors who have piloted these assessments in their classrooms and the students who have taken them. Finally, a special thank you goes to Santiago Bohórquez Correa at Universidad EAFIT for translating our economic statistics skills assessment (ESSA) into Spanish.

References

Adams, W.K., and C.E. Wieman. 2011. Development and Validation of Instruments to Measure Learning of Expert-Like Thinking. *International Journal of Science Education* 33(9): 1289–1312. https://doi.org/10.1080/09500693.2010.512369.

Becker, W.E., and J.R. Powers. 2001. Student performance, attrition, and class size given missing student data. *Economics of Education Review* 20(4): 377–388.

Bottan, D., D. McKee, G. Orlov, and A. McDougall. 2022. Racial and gender achievement gaps in an economics classroom. *International Review of Economics Education* 40: 100239. https://doi.org/10.1016/j.iree.2022.100239.

Coates, D., B.R. Humphreys, J. Kane, and M.A. Vachris. 2004. “No significant distance” between face-to-face and online instruction: Evidence from principles of economics. *Economics of Education Review* 23(5): 533–546.

¹ In the case of smaller class sizes, the available information on subgroups of students may be limited in order to protect student identities. Specifically, the system will not show statistics for subgroups with fewer than ten students.
Cornell Center for Teaching Innovation. 2022. Active Learning Initiative. https://teaching.cornell.edu/programs/faculty-instructors/active-learning-initiative

Crouch, C.H., and E. Mazur. 2001. Peer Instruction: Ten years of experience and results. *American Journal of Physics* 69(9): 970–977. https://doi.org/10.1119/1.1374249.

Emerson, T. L., and B. A. Taylor. 2004. Comparing student achievement across experimental and lecture-oriented sections of a principles of microeconomics course. *Southern Economic Journal*, 672–693.

Freeman, S., S.L. Eddy, M. McDonough, M.K. Smith, N. Okoroafor, H. Jordt, and M.P. Wenderoth. 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences* 111(23): 8410–8415. https://doi.org/10.1073/pnas.1319030111.

McKee, D., and G. Orlov. 2022. The Economics Statistics Skills Assessment (ESSA). Working paper.

Orlov, G., D. McKee, J. Berry, A. Boyle, T. DiCiccio, T. Ransom, et al. 2021a. Learning during the COVID-19 pandemic: It is not who you teach, but how you teach. *Economics Letters* 202: 109812. https://doi.org/10.1016/j.econlet.2021.109812.

Orlov, G., D. McKee, I. Foster, D. Bottan, S. Thomas, and J. Wissink. 2021b. Identifying Students-at-Risk Using a New Math Skills Assessment. *American Economic Review Papers & Proceedings*.

Rawson, K.A., and J. Dunlosky. 2012. When Is Practice Testing Most Effective for Improving the Durability and Efficiency of Student Learning? *Educational Psychology Review* 24(3): 419–435. https://doi.org/10.1007/s10648-012-9203-1.

Salehi, S., S. Cotner, S.M. Azarin, E.E. Carlson, M. Driessen, and V.E. Ferry, et al. 2019. Gender Performance Gaps Across Different Assessment Methods and the Underlying Mechanisms: The Case of Incoming Preparation and Test Anxiety. *Frontiers in Education* 4. https://doi.org/10.3389/feduc.2019.00107.

Wainer, H., and D. Thissen. 1993. Combining Multiple-Choice and Constructed-Response Test Scores: Toward a Marxist Theory of Test Construction. *Applied Measurement in Education* 6(2): 103–118. https://doi.org/10.1207/s15324818ame0602_1.

Walstad, W.B. 1998. Multiple choice tests for the economics course. *Teaching Undergraduate Economics: A Handbook for Instructors* 287: 304.

Walstad, W.B., and W.E. Becker. 1994. Achievement Differences on Multiple-Choice and Essay Tests in Economics. *The American Economic Review* 84(2): 193–196.

Walstad, W.B., and K. Rebeck. 2008. The Test of Understanding of College Economics. *American Economic Review* 98(2): 547–551. https://doi.org/10.1257/aer.98.2.547.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.