Copyright © 2026 by Cherkas Global University
All rights reserved.
Published in the USA

European Journal of Contemporary Education
E-ISSN 2305-6746
2026. 15(2): 326-337
DOI: 10.13187/ejced.2026.2.326
https://ejce.cherkasgu.press

IMPORTANT NOTICE! Any copying, reproduction, distribution, republication (in whole or in part), or otherwise commercial use of this work in violation of the author's rights will be prosecuted in accordance with international law. The use of hyperlinks to the work will not be considered copyright infringement.




# How Case-Based Learning Shapes Economic Reasoning in Undergraduate Students

Sergey A. Tronin [a, *]

[a] Financial University under the Government of the Russian Federation, Moscow, Russian Federation

**Abstract**

Despite the growing emphasis on active pedagogies in higher education, the dominant instructional mode in undergraduate economics remains the traditional lecture. This study investigated whether a semester-long case-based learning (CBL) intervention could improve economic reasoning among second-year undergraduate students enrolled in microeconomics courses at a Polish university. A quasi-experimental pre-test–post-test design with non-equivalent control groups was employed. The experimental group (n = 72) participated in structured CBL sessions built around authentic economic policy scenarios, whereas the control group (n = 68) followed a conventional lecture-and-textbook format covering identical content. Economic reasoning was assessed through an adapted version of the Test of Understanding in College Economics (TUCE-4) and a purpose-built Economic Reasoning Assessment (ERA) rubric aligned with the revised Bloom's taxonomy. Results indicated that students in the CBL group demonstrated statistically significant gains in overall TUCE scores ($t = 3.41$, $p < .001$, Cohen's $d = 0.62$) and in higher-order reasoning categories – analysis, evaluation, and application – compared with the control group. A chi-squared analysis confirmed a significantly higher proportion of CBL students reaching the "proficient" threshold on the ERA ($\chi^2 = 9.74$, $df = 1$, $p = .002$). No significant between-group differences were observed for lower-order cognitive categories (recognition and comprehension). Qualitative data from semi-structured interviews with a purposive subsample (n = 24) revealed that students attributed their reasoning gains to the necessity of confronting ambiguity in real-world cases and to the collaborative argumentation that case discussions fostered. These findings suggest that CBL is a viable complement to traditional instruction for cultivating the analytical and evaluative competencies that economics curricula aspire to develop.

**Keywords:** case-based learning, economic reasoning, undergraduate economics education, active learning, Bloom's taxonomy, TUCE, quasi-experimental design.

* Corresponding author
E-mail addresses: tron1977@rambler.ru (S.A. Tronin)

## 1. Introduction

A persistent tension runs through the scholarship on economics education. On the one hand, disciplinary bodies and employers expect graduates to "think like economists" – to identify trade-offs, reason about incentives, evaluate competing policy alternatives, and construct arguments from evidence (Siegfried et al., 1991; Becker, Watts, 2001). On the other hand, the modal classroom experience in undergraduate economics worldwide continues to be the lecture, a format whose effectiveness in nurturing higher-order cognitive skills has been questioned for decades (Becker, Watts, 1996; Watts, Schaur, 2011). Surveys of economics departments in the United States, the United Kingdom, and continental Europe consistently report that chalk-and-talk remains the primary instructional strategy, with interactive and experiential methods occupying a marginal role (Watts, Schaur, 2011; Goffe, Kauper, 2014).

The call for pedagogical renewal in economics is not merely an abstract aspiration. Employers in finance, consulting, and the public sector increasingly emphasize analytical reasoning, communication, and problem-solving – competencies that align with the upper tiers of Bloom's revised taxonomy (Anderson, Krathwohl, 2001) but that conventional lecture delivery does not systematically cultivate (Becker, 1997). The gap between intended learning outcomes and actual instructional practice has prompted a growing interest in active-learning techniques, including classroom experiments, simulations, flipped formats, and case-based learning (CBL) (Simkins, Maier, 2010; Allgood et al., 2015).

Among these alternatives, CBL occupies a distinctive niche. Unlike laboratory experiments or simulations, which often isolate a single market mechanism, well-designed cases immerse students in the messy complexity of real-world economic problems – problems in which data are incomplete, multiple theoretical frameworks compete for explanatory relevance, and policy recommendations carry distributional consequences that extend beyond textbook equilibria (Carlson, Schodt, 1995; Velenchik, 1995). Precisely because cases resist neat solutions, they compel students to exercise the analytical, evaluative, and synthetic reasoning that Bloom's taxonomy positions at the apex of cognitive development (Volpe, 2015).

Despite the pedagogical appeal of CBL, empirical evidence on its effectiveness in economics courses remains limited and methodologically uneven. The landmark review by Volpe (2015) concludes that research is "generally mixed but broadly supports the view that [CBL] facilitates the retention of information and helps the acquisition of a deeper understanding of the subject." Ray (2018) provides one of the few controlled comparisons, reporting statistically significant improvements in final examination scores after introducing cases into two elective economics courses. Carlson and Velenchik (2006) offer practical guidance on case implementation but do not present experimental data. More broadly, the literature on active learning in economics has tended to rely on student satisfaction surveys and course-grade comparisons rather than on instruments designed to capture the quality of economic reasoning (Watts, Schaur, 2011). There is therefore a gap in the literature: a need for quasi-experimental studies that employ validated cognitive assessments and that examine not merely whether CBL raises test scores but whether it differentially affects lower-order and higher-order dimensions of economic reasoning.

The present study addresses this gap. Its primary objective was to evaluate the impact of a structured CBL intervention on the economic reasoning of second-year undergraduate students enrolled in microeconomics courses at a large Polish university. The study pursued three research questions:

RQ1. Do students who participate in a semester-long CBL program demonstrate greater overall gains in economic understanding, as measured by the TUCE, than students who follow a conventional lecture format?

RQ2. Does CBL differentially improve higher-order reasoning (application, analysis, evaluation) compared with lower-order cognition (recognition, comprehension)?

RQ3. How do students themselves perceive the contribution of CBL to their economic reasoning?

A quasi-experimental pre-test–post-test design with non-equivalent control groups was used. The experimental group received CBL sessions integrated into regular microeconomics instruction; the control group covered the same syllabus through lectures and textbook exercises. Economic reasoning was assessed through two instruments: an adapted subset of items from the TUCE-4 (Walstad, Rebeck, 2008) and a purpose-designed Economic Reasoning Assessment rubric that mapped

student responses onto the cognitive levels of Anderson and Krathwohl's (2001) revised taxonomy. Semi-structured interviews with a purposive subsample supplemented the quantitative findings.

The study contributes to the literature in three ways. First, it provides quasi-experimental evidence from a Central European institutional context, where research on active pedagogy in economics is sparse. Second, it disaggregates cognitive outcomes by Bloom's taxonomy level, thereby addressing the common criticism that aggregate test-score comparisons obscure the specific reasoning processes that CBL is theorized to develop. Third, it triangulates quantitative and qualitative data to offer a richer account of how students experience– and benefit from – case-based instruction in economics.

## 2. Literature Review

### 2.1. The Case Method in Economics Education

The case method has a long pedigree in professional education, particularly in law and business (Christensen, Hansen, 1987). Its migration into undergraduate economics courses, however, has been slower and more contested. Carlson and Schodt (1995) were among the earliest advocates, arguing that case teaching moves students "beyond the lecture" by requiring them to apply theoretical concepts to situations that do not map cleanly onto textbook models. Velenchik (1995) identified four pedagogical benefits of cases in economics: they motivate the study of theory by making its relevance visible; they require students to select and apply appropriate analytical tools; they expose the institutional and political context in which economic decisions are made; and they develop collaborative skills through structured discussion. Buckles (1998) reinforced these arguments, positioning CBL as an "effective active learning technique" particularly suited to courses in which policy analysis is a stated learning outcome. More recently, Volpe (2015) provided a comprehensive review that traced the diffusion of case teaching from elite business schools to liberal arts colleges and European universities. His survey of the empirical literature concluded that the evidence, although not unequivocal, favoured CBL over purely lecture-based instruction in terms of student engagement, retention, and the development of higher-order skills. Popescu (2014) similarly emphasized that the distinguishing feature of an economics teaching case is that it provides information but neither analysis nor conclusions; the analytical work is performed by students during classroom discussion.

Ray (2018) conducted one of the few quantitative assessments of CBL in economics. In a study at a small liberal-arts college, she introduced case modules into two elective courses and compared outcomes with sections taught by the same instructor without cases. Both a pooled OLS regression and a difference-in-differences analysis yielded statistically significant positive effects of the case treatment on final-examination performance. Students in the CBL sections also ranked "learning how to use economics to solve real problems" as the most valuable aspect of the course – an observation consistent with the motivation thesis advanced by Velenchik (1995).

### 2.2. Active Learning and Economic Reasoning

The broader literature on active learning in economics provides useful context for the present study. Simkins and Maier (2010) offered a systematic framework – "Just-in-Time Teaching" (JiTT) – that blends pre-class preparation with interactive class activities. Their evidence suggested that JiTT improves both student engagement and conceptual understanding. Emerson and Taylor (2004) reported gains from classroom experiments, using the TUCE as the outcome measure. Goffe and Kauper (2014) examined the persistence of the lecture format and identified structural factors – class size, disciplinary culture, and departmental incentives – that inhibit the adoption of interactive methods even when evidence of their superiority is available.

A critical insight from this literature is that not all active-learning techniques affect all dimensions of cognition equally. Freeman et al. (2014), in a meta-analysis spanning STEM disciplines, demonstrated that active learning reduces failure rates and raises average examination scores; however, the meta-analysis did not disaggregate outcomes by cognitive level. Within economics specifically, Buckles and Siegfried (2006) argued that well-designed multiple-choice items can capture some higher-order skills but that the field has underinvested in assessments that separately measure recognition, comprehension, application, analysis, and evaluation. This observation motivates the present study's decision to assess economic reasoning at each level of Bloom's taxonomy.

### 2.3. Bloom's Taxonomy and the Assessment of Economic Cognition

Bloom's original taxonomy (Bloom, 1956) has been a foundational framework in educational measurement. The revised version by Anderson and Krathwohl (2001) reorganized the six

cognitive levels – Remember, Understand, Apply, Analyze, Evaluate, Create – and introduced a two-dimensional structure that crosses cognitive processes with knowledge types. In economics education, the TUCE has historically classified items into three cognitive categories: recognition and understanding, explicit application, and implicit application (Walstad et al., 2007; Walstad, Rebeck, 2008). Although the TUCE's taxonomy is broadly compatible with Bloom's framework, it does not fully map onto the six revised levels, particularly at the evaluative and creative ends of the continuum.

Several scholars have argued that economics education must aim higher on the Bloom hierarchy. Becker (1997) observed that the predominance of multiple-choice examinations in introductory courses implicitly privileges recognition and comprehension over deeper reasoning. Siegfried and Walstad (1998) noted that gains on the TUCE correlate only weakly with the ability to transfer economic reasoning to novel contexts – a skill that resides at the application and analysis levels. More recently, Allgood et al. (2015) proposed that the signature skill of an economics graduate – the capacity to "think at the margin" in unfamiliar situations – requires instructional strategies that deliberately scaffold higher-order cognition. CBL, with its emphasis on ambiguity, trade-offs, and evidence-based argumentation, is a natural candidate for such scaffolding.

### 2.4. Gaps in the Literature

Three gaps emerge from the foregoing review. First, most studies of CBL in economics are small-scale, single-course designs that do not control for instructor effects or prior student ability. Second, outcome measures tend to be course grades or student-satisfaction surveys rather than validated cognitive assessments that distinguish among reasoning levels. Third, the literature is geographically concentrated in the United States and the United Kingdom; evidence from Central and Eastern European higher education, where the lecture tradition is particularly entrenched, is scarce. The present study was designed to address all three limitations.

## 3. Methods

### 3.1. Research Design

A quasi-experimental pre-test–post-test design with non-equivalent control groups was employed. Random assignment at the individual level was not feasible because students were enrolled in intact course sections. To mitigate selection bias, the study used sections taught by two instructors of comparable rank and experience who followed an identical syllabus and shared a common item bank for examinations. One instructor's two sections (combined n = 72) received the CBL treatment; the other instructor's two sections (combined n = 68) served as the control.

### 3.2. Participants

Participants were 140 second-year undergraduate students pursuing a Bachelor's degree in Economics at a large public university in southern Poland during the spring semester of the 2023–2024 academic year. The experimental group comprised 72 students (41 female, 31 male; mean age 20.8 years, SD = 0.9). The control group comprised 68 students (38 female, 30 male; mean age 20.6 years, SD = 1.1). All participants had completed the same first-year sequence in introductory economics and quantitative methods, and all were concurrently enrolled in the required second-year Microeconomics course. Students were informed of the study and provided written consent; participation in the assessment activities was voluntary. The study was approved by the university's ethical review board.

### 3.3. Intervention: The CBL Program

The CBL program consisted of 10 structured case sessions integrated into the 14-week semester. Each session lasted 90 minutes and replaced one of the two weekly contact hours that the experimental sections would otherwise have spent on conventional problem sets. The remaining weekly contact hour continued as a standard lecture, ensuring that theoretical content coverage was equivalent across conditions.

Cases were drawn from publicly available sources – journalistic accounts, policy reports, statistical databases – and adapted by the research team to align with the microeconomics syllabus. Each case centred on a real economic event or policy dilemma relevant to the European context. Illustrative examples included: the impact of rent-control policies in Berlin and Vienna; the competitive dynamics of low-cost airlines in the EU single market; price-discrimination strategies in pharmaceutical markets; and the economic consequences of the EU carbon border adjustment mechanism. Cases were distributed to students 48 hours before the session to allow independent preparation. Classroom sessions followed a structured protocol adapted from the

Harvard Business School case-teaching model (Christensen, Hansen, 1987) and modified to suit the economics context along lines recommended by Carlson and Velenchik (2006). Each session comprised three phases: (1) an opening phase (15 minutes), in which the instructor posed a framing question and solicited initial reactions; (2) an analysis phase (50 minutes), in which small groups of four to five students worked through a set of guiding questions that required them to identify the relevant economic concepts, marshal evidence from the case materials, construct and compare alternative policy recommendations, and anticipate counterarguments; and (3) a synthesis phase (25 minutes), in which groups reported their analyses, the instructor facilitated a whole-class debate, and key takeaways were summarized on the board.

### 3.4. Control Condition

The control group covered the same syllabus topics in the same sequence but exclusively through lectures and textbook-based problem sets. Problem sets contained end-of-chapter exercises and numerical problems drawn from the course textbook. Students worked on these individually outside of class and reviewed solutions during a weekly tutorial session. No case materials were used in the control sections.

### 3.5. Instruments

**Test of Understanding in College Economics (TUCE-4), adapted.** The TUCE-4 micro-economics form (Walstad et al., 2007; Walstad, Rebeck, 2008) is a 30-item multiple-choice instrument that has been widely used in economics education research. For the present study, 20 items were selected to match the content coverage of the second-year microeconomics syllabus (consumer theory, production and costs, market structures, factor markets, and the microeconomic role of government). Items were translated into Polish by two bilingual economists and back-translated by an independent translator; discrepancies were resolved by consensus. The adapted instrument was pilot-tested with 35 third-year students who were not part of the study sample; Cronbach's alpha for the pilot was 0.78, which is consistent with the reliability range (0.72–0.82) reported for the original TUCE-4 (Walstad et al., 2007). The 20 items spanned the TUCE's three cognitive categories: recognition and understanding (7 items), explicit application (7 items), and implicit application (6 items).

**Economic Reasoning Assessment (ERA).** To capture reasoning dimensions that multiple-choice items cannot fully access, the research team developed a four-item open-ended assessment. Each item presented a brief economic scenario and posed a question that required students to: (a) identify the core economic problem, (b) apply at least one theoretical framework, (c) evaluate trade-offs inherent in two or more policy options, and (d) justify a recommendation with reference to evidence. Responses were scored on a 16-point rubric (four items × four points each) that mapped onto four cognitive levels drawn from Anderson and Krathwohl's (2001) revised taxonomy: Understanding (1 point – correctly identifies the economic problem), Application (2 points – applies a relevant theoretical model), Analysis (3 points – decomposes the problem and evaluates trade-offs), and Evaluation (4 points – formulates a justified policy recommendation that addresses counterarguments). Two trained raters scored all responses independently; inter-rater reliability was assessed using Cohen's kappa ($\kappa = 0.83$), indicating substantial agreement (Landis, Koch, 1977). Discrepancies of more than one point on any item were resolved through discussion.

**Semi-structured interviews.** A purposive subsample of 24 students from the experimental group (12 female, 12 male) was selected to represent a range of pre-test performance levels (high, medium, low). Interviews lasted 25–40 minutes and were conducted in Polish within two weeks of the post-test. The interview protocol included seven open-ended questions exploring students' perceptions of how CBL affected their understanding of economic concepts, their reasoning processes during case analysis, and any challenges they encountered. Interviews were audio-recorded, transcribed verbatim, and analysed thematically following the six-phase procedure of Braun and Clarke (2006).

### 3.6. Procedure

The pre-test (adapted TUCE and ERA) was administered during the first week of the semester, before any treatment had begun. All students completed the same instruments under identical conditions (60 minutes, invigilated classroom setting). The CBL intervention then ran across weeks 2–13. The post-test was administered during week 14, using parallel forms of both instruments to minimise test-retest effects: TUCE items were identical (as is standard in pre-post TUCE designs; Walstad, Rebeck, 2008), while ERA scenarios were different in surface features but structurally isomorphic to the pre-test scenarios. Interviews were conducted during weeks 15–16.

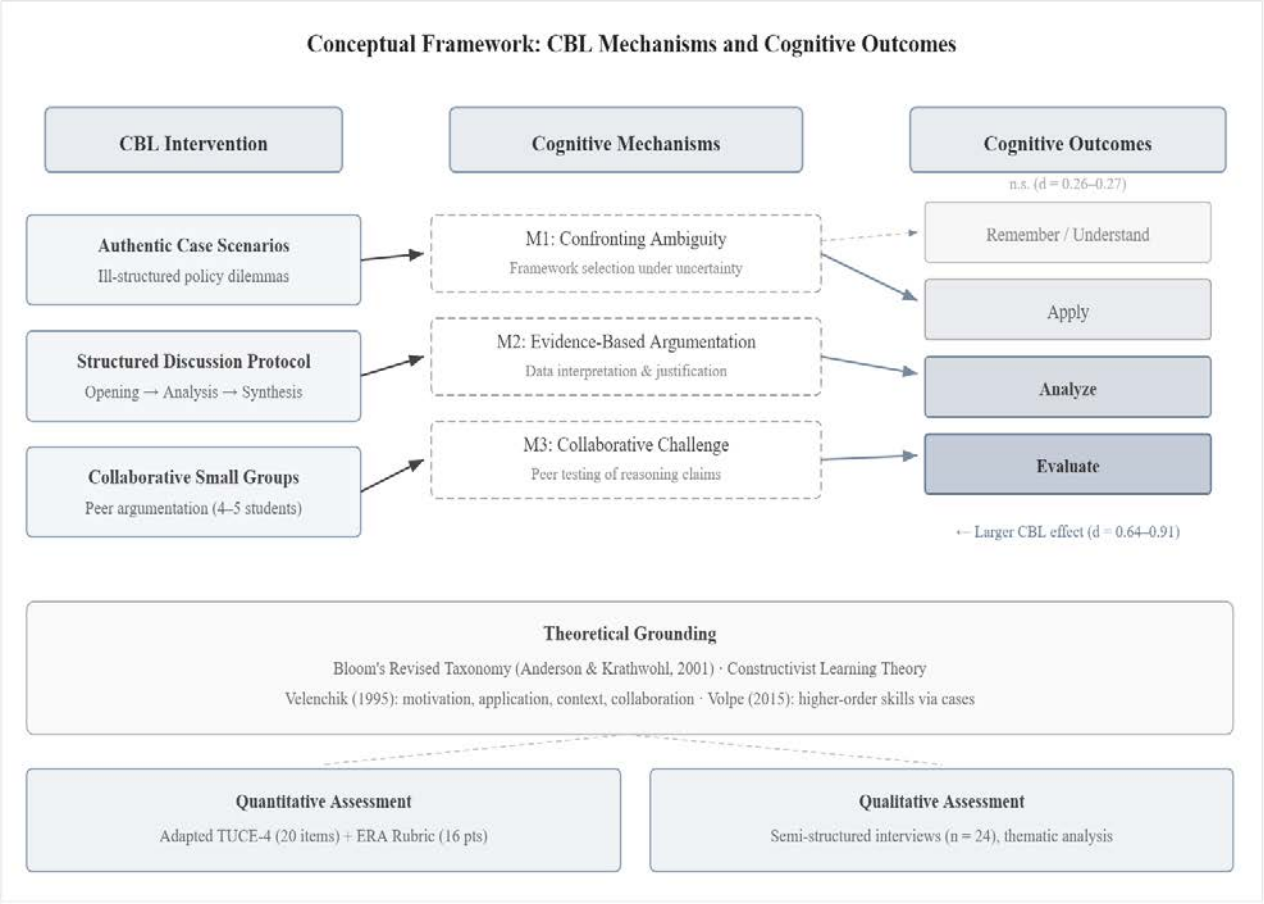



**Fig. 1.** Conceptual framework linking case-based learning intervention components, hypothesised cognitive mechanisms, and cognitive outcomes mapped onto the revised Bloom's taxonomy. Solid arrows indicate the theorised primary pathways; dashed arrows indicate secondary pathways. Effect-size annotations reflect observed between-group differences (see Section 4)

The conceptual framework (Figure 1) maps the theorised causal chain from the three structural components of the CBL intervention – authentic case scenarios, the structured three-phase discussion protocol, and collaborative small-group work – through three cognitive mechanisms identified in the qualitative data (confronting ambiguity, evidence-based argumentation, and collaborative challenge) to the higher-order levels of the revised Bloom's taxonomy (Anderson, Krathwohl, 2001). Solid arrows denote primary pathways supported by both quantitative and qualitative evidence; the dashed arrow indicates a secondary link between ambiguity exposure and the application level that was supported only by the TUCE data. The shading gradient in the Cognitive Outcomes column reflects the observed pattern of effect sizes: CBL effects were negligible at the Remember/Understand level (d = 0.26–0.27, n.s.) and progressively larger at the Apply (d = 0.64–0.67), Analyse (d = 0.91), and Evaluate (d = 0.80) levels.

### 3.7. Data Analysis

Quantitative analyses were performed using IBM SPSS Statistics (version 28). Prior to parametric testing, the distributional properties of all continuous variables were examined. Skewness and kurtosis values for pre-test and post-test TUCE scores ranged from −0.38 to 0.41 and from −0.52 to 0.29, respectively; for ERA scores, skewness ranged from −0.27 to 0.34 and kurtosis from −0.61 to 0.18. All values fell within the conventional ±1.0 threshold recommended by George and Mallery (2019), supporting the assumption of approximate normality. Pre-test scores were compared across groups using independent-samples t-tests to check for baseline equivalence. Post-test comparisons employed independent-samples t-tests and, where appropriate, analysis of covariance (ANCOVA) with pre-test scores as the covariate to adjust for any residual baseline differences. Before conducting ANCOVA, the homogeneity-of-regression-slopes assumption was verified by testing the Group × Pre-test interaction term, which was non-significant for both the TUCE ($F(1, 136) = 0.42$, $p = .518$) and the ERA ($F(1, 136) = 0.61$, $p = .436$). Levene's test confirmed

homogeneity of variances for post-test TUCE scores ($F(1, 138) = 1.03$, $p = .312$) and post-test ERA scores ($F(1, 138) = 1.47$, $p = .227$). Effect sizes were reported as Cohen's d. For the ERA, the proportion of students reaching the "proficient" threshold (≥ 12 out of 16 points) was compared using Pearson's chi-squared test. The significance level was set at $\alpha = .05$. Qualitative data were analysed using reflexive thematic analysis (Braun, Clarke, 2006).

## 4. Results

### 4.1. Baseline Equivalence

Pre-test scores confirmed that the two groups were comparable at baseline. On the adapted TUCE, the experimental group scored M = 8.47 (SD = 2.84) and the control group scored M = 8.21 (SD = 2.71) out of 20. The difference was not statistically significant ($t(138) = 0.56$, $p = .576$, $d = 0.10$). On the ERA, the experimental group scored M = 5.92 (SD = 2.19) and the control group scored M = 5.74 (SD = 2.30) out of 16 ($t(138) = 0.48$, $p = .631$, $d = 0.08$). The groups also did not differ significantly in gender composition ($\chi^2(1) = 0.03$, $p = .862$) or cumulative GPA from the first year ($t(138) = 0.71$, $p = .479$). These results support the assumption of baseline equivalence.

### 4.2. TUCE Outcomes

Post-test TUCE results are presented in Table 1.

**Table 1.** Pre-test and Post-test TUCE Scores by Group

| Measure | Group | n | Pre-test M (SD) | Post-test M (SD) | Gain M (SD) | t | p | d |
|---|---|---|---|---|---|---|---|---|
| TUCE Total (0–20) | CBL | 72 | 8.47 (2.84) | 14.19 (2.92) | 5.72 (2.46) | – | – | – |
| | Control | 68 | 8.21 (2.71) | 12.07 (3.18) | 3.86 (2.63) | – | – | – |
| Between-group post-test | | | | | | 4.14 | <.001 | 0.69 |
| Between-group gain | | | | | | 3.41 | <.001 | 0.62 |

The CBL group achieved a mean post-test score of 14.19 (SD = 2.92), compared with 12.07 (SD = 3.18) for the control group. The between-group difference in post-test scores was statistically significant ($t(138) = 4.14$, $p < .001$, $d = 0.69$). When pre-test scores were entered as a covariate in an ANCOVA, the adjusted post-test means remained significantly different ($F(1, 137) = 16.82$, $p < .001$, partial $\eta^2 = .109$).

Analysis by cognitive category revealed a nuanced pattern (Table 2).

**Table 2.** Post-test TUCE Scores by Cognitive Category

| Cognitive Category | Items | CBL M (SD) | Control M (SD) | t(138) | p | d |
|---|---|---|---|---|---|---|
| Recognition and Understanding | 7 | 5.39 (1.22) | 5.06 (1.34) | 1.54 | .126 | 0.26 |
| Explicit Application | 7 | 4.92 (1.31) | 3.93 (1.42) | 4.27 | <.001 | 0.72 |
| Implicit Application | 6 | 3.88 (1.18) | 3.09 (1.29) | 3.78 | <.001 | 0.64 |

No statistically significant difference was observed in recognition and understanding items ($t(138) = 1.54$, $p = .126$, $d = 0.26$). By contrast, the CBL group substantially outperformed the control group on explicit application items ($t(138) = 4.27$, $p < .001$, $d = 0.72$) and implicit application items ($t(138) = 3.78$, $p < .001$, $d = 0.64$). The pattern supports the hypothesis that CBL differentially enhances higher-order economic reasoning.

### 4.3. ERA Outcomes

Table 3 presents pre-test and post-test ERA results.

**Table 3.** Pre-test and Post-test ERA Scores by Group

| Measure | Group | n | Pre-test M (SD) | Post-test M (SD) | Gain M (SD) |
|---|---|---|---|---|---|
| ERA Total (0–16) | CBL | 72 | 5.92 (2.19) | 11.43 (2.38) | 5.51 (2.11) |
| | Control | 68 | 5.74 (2.30) | 9.18 (2.72) | 3.44 (2.34) |
| Between-group post-test | | | | $t(138) = 5.29$, $p < .001$, $d = 0.88$ | |
| Between-group gain | | | | $t(138) = 5.05$, $p < .001$, $d = 0.93$ | |

The mean post-test ERA score in the CBL group ($M = 11.43$, $SD = 2.38$) exceeded that of the control group ($M = 9.18$, $SD = 2.72$) by 2.25 points, a difference that was highly significant ($t(138) = 5.29$, $p < .001$) with a large effect size ($d = 0.88$). The gain score comparison was similarly significant ($t(138) = 5.05$, $p < .001$, $d = 0.93$).

Disaggregation by rubric level provided further insight (Table 4).

**Table 4.** Mean ERA Score by Rubric Level (Post-test)

| Rubric Level | Maximum | CBL M (SD) | Control M (SD) | t(138) | p | d |
|---|---|---|---|---|---|---|
| Understanding | 4 | 3.42 (0.64) | 3.24 (0.71) | 1.57 | .118 | 0.27 |
| Application | 4 | 3.08 (0.78) | 2.53 (0.87) | 3.93 | <.001 | 0.67 |
| Analysis | 4 | 2.76 (0.82) | 1.97 (0.91) | 5.36 | <.001 | 0.91 |
| Evaluation | 4 | 2.17 (0.94) | 1.44 (0.88) | 4.71 | <.001 | 0.80 |

Consistent with the TUCE results, the largest between-group differences appeared at the analysis ($d = 0.91$) and evaluation ($d = 0.80$) levels, while understanding yielded only a non-significant difference ($d = 0.27$). The application level fell in between ($d = 0.67$).

When students were classified as "proficient" (ERA ≥ 12) or "not proficient" (ERA < 12), a markedly higher proportion of CBL students met the threshold (Table 5).

**Table 5.** Proportion of Students Reaching ERA Proficiency Threshold (≥ 12/16)

| Group | n | Proficient n (%) | Not Proficient n (%) | $\chi^2$ | p |
|---|---|---|---|---|---|
| CBL | 72 | 34 (47.2) | 38 (52.8) | | |
| Control | 68 | 16 (23.5) | 52 (76.5) | | |
| Total | 140 | 50 (35.7) | 90 (64.3) | 9.74 | .002 |

The chi-squared test confirmed that the association between group membership and proficiency status was statistically significant ($\chi^2(1) = 9.74$, $p = .002$). Students in the CBL group were roughly twice as likely to reach the proficiency threshold as their counterparts in the control group.

Figure 2 presents a disaggregated comparison of post-test performance across cognitive levels. Panel A displays normalised TUCE-4 means for the three cognitive categories defined by Walstad et al. (2007); Panel B displays ERA rubric means for four levels aligned with Anderson and Krathwohl's (2001) taxonomy. Across both instruments, the between-group gap widens monotonically as one ascends the cognitive hierarchy, confirming the central hypothesis that CBL differentially strengthens higher-order economic reasoning.

The largest effect sizes were observed for the Analysis rubric level ($d = 0.91$) and the Explicit Application TUCE category ($d = 0.72$). By contrast, neither the Recognition and Understanding TUCE category nor the Understanding ERA level yielded a statistically significant between-group difference, suggesting that conventional lectures are adequate for cultivating foundational factual knowledge but insufficient for developing the evaluative and analytical competencies that economics programmes aspire to instill.

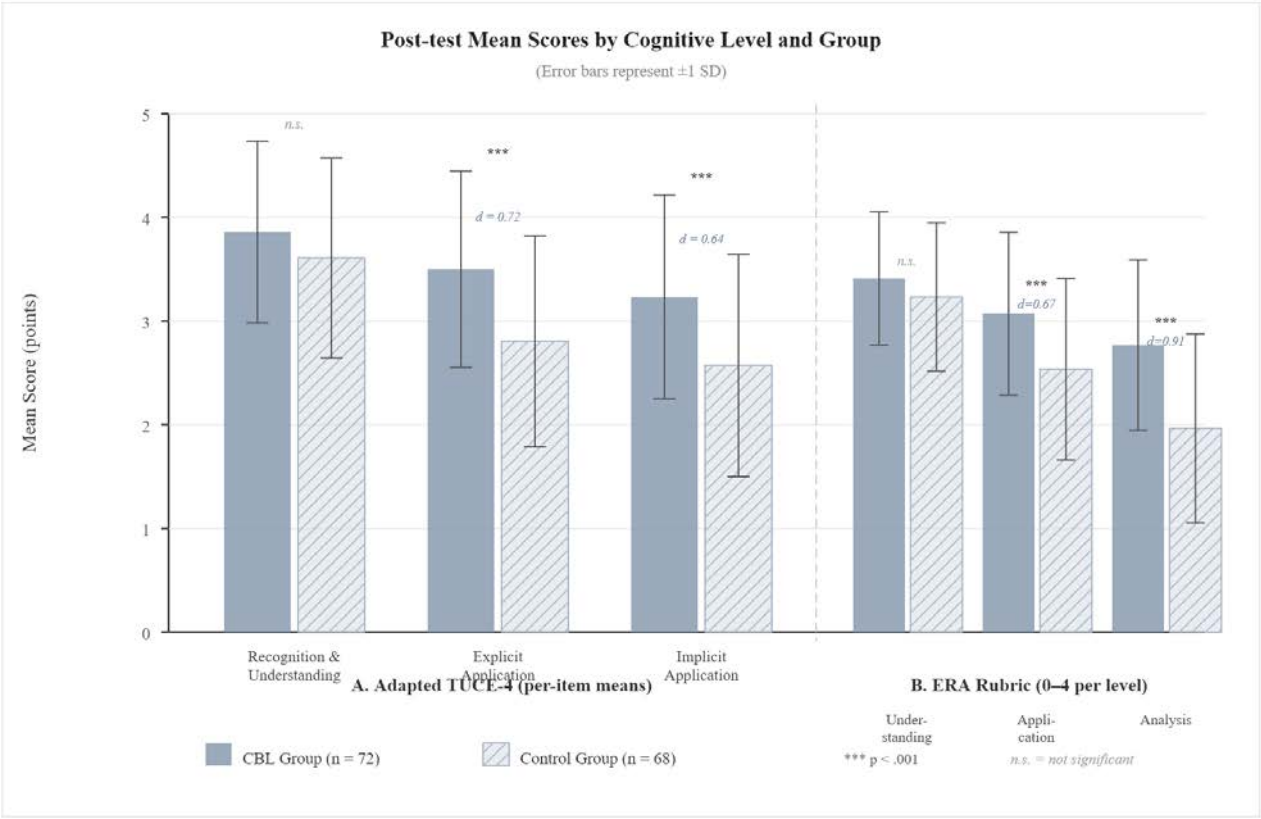



**Fig. 2.** Post-test mean scores by cognitive level for the CBL and control groups. Panel A displays normalised per-item means (×5) on the three cognitive categories of the adapted TUCE-4. Panel B displays mean scores on the four rubric levels of the Economic Reasoning Assessment (ERA; 0–4 scale). Error bars represent ±1 SD. Significance markers: *** p < .001; n.s. = not significant. Effect sizes (Cohen's d) are reported for significant comparisons

### 4.4. Qualitative Findings

Thematic analysis of the 24 interviews yielded four themes.

**Theme 1: Confronting ambiguity.** Students repeatedly emphasised that cases forced them to work without clear-cut answers. A representative comment was: "In a normal class, the professor gives you the model and you plug in numbers. In the case, nobody tells you which model to use. You have to figure it out, and sometimes two models give different answers" (Participant 7, female, high pre-test). This perception aligns with the theorised mechanism by which CBL cultivates analytical reasoning: the absence of a single correct solution compels students to evaluate the comparative explanatory power of competing frameworks.

**Theme 2: Learning to argue with evidence.** Fifteen of the 24 interviewees mentioned that preparing for case discussions required them to locate and interpret data-price indices, market-share statistics, regulatory documents – that they would not have encountered in a textbook exercise. Several students noted that the experience of defending a policy recommendation in front of peers improved their capacity to construct evidence-based arguments: "You can't just say 'it's inefficient.' Someone will ask, 'compared to what?' and 'show me the data.' That makes you think more carefully" (Participant 12, male, medium pre-test).

**Theme 3: Collaborative reasoning.** Students described the small-group phase of case sessions as particularly productive. Working in groups of four or five, they reported that disagreements within the group – about which assumptions were reasonable, which externalities were most significant, or which stakeholder perspective deserved priority – stimulated deeper thinking: "When someone disagrees with you, you have to think about why you believe what you believe. That's harder than solving a textbook problem, but I learned more from it" (Participant 19, female, low pre-test).

**Theme 4: Transfer to other contexts.** Several students spontaneously described applying case-derived reasoning to situations outside the microeconomics course – in discussions of macroeconomic policy, in a concurrent business-ethics elective, or in everyday conversations

about economic news: "After the airline case, I started reading news about competition differently. I actually think about market structure now, not just prices" (Participant 3, male, high pre-test). While anecdotal, these observations suggest that CBL may promote the kind of conceptual transfer that Allgood et al. (2015) identify as a hallmark of genuine economic reasoning.

## 5. Discussion

### 5.1. CBL and Overall Economic Understanding

The first research question asked whether CBL students would demonstrate greater overall gains in economic understanding. The results provide an affirmative answer. The CBL group outperformed the control group by approximately 2.1 points on the 20-item adapted TUCE, a difference that corresponds to a medium-to-large effect (d = 0.69). This finding is broadly consistent with the positive effects reported by Ray (2018) and with the meta-analytic evidence that active learning outperforms traditional lecturing across STEM disciplines (Freeman et al., 2014). The adjusted analysis (ANCOVA) confirmed that the difference was not attributable to pre-existing group differences. It is worth noting, however, that the control group also registered meaningful gains (M gain = 3.86 points), reflecting the fact that conventional instruction is not without value. The question is not whether lectures "work" but whether supplementing them with structured case analysis adds value at the higher end of the cognitive spectrum – a question addressed by the second research question.

### 5.2. Differential Effects Across Cognitive Levels

The most striking finding of the study is the pattern of differential effects across cognitive categories. On both the TUCE and the ERA, between-group differences were small and non-significant for lower-order cognition (recognition, understanding) but large and significant for higher-order cognition (application, analysis, evaluation). This pattern is consistent with the theoretical prediction that CBL's principal mechanism of action is the scaffolding of higher-order reasoning through engagement with ill-structured problems (Velenchik, 1995; Volpe, 2015). Students in both groups were exposed to the same theoretical content through lectures; the CBL group additionally had the opportunity to exercise that content in contexts that demanded selection, integration, and critical evaluation – activities that reside at the upper levels of Bloom's taxonomy.

The finding has practical implications. If the goal of an introductory or intermediate microeconomics course is limited to ensuring that students can recognise definitions and recall standard models, then CBL offers little marginal benefit over well-delivered lectures. If, however, the goal includes the development of analytical and evaluative reasoning – "thinking like an economist" in its fullest sense – then CBL appears to provide a substantial and measurable advantage. Given that most economics programmes list higher-order reasoning among their intended learning outcomes (Siegfried et al., 1991; Allgood et al., 2015), the implication is that CBL deserves a more prominent place in the curriculum than it currently occupies.

### 5.3. Student Perceptions and Mechanisms

The qualitative data shed light on the mechanisms through which CBL may exert its effects. Three mechanisms emerged with particular clarity. First, the ambiguity inherent in authentic cases pushes students beyond the "plug-and-chug" mentality that textbook exercises reinforce. When no single model provides a complete answer, students must compare frameworks – a process that maps onto the "analyse" and "evaluate" levels of Bloom's taxonomy. Second, the requirement to argue with evidence – to justify claims not merely with theoretical logic but with data drawn from the case materials – develops a skill set that is central to economic reasoning yet rarely practiced in lecture-based courses. Third, collaborative argumentation in small groups creates a social context in which individual reasoning is tested, challenged, and refined. These three mechanisms – confronting ambiguity, arguing with evidence, and collaborative challenge – collectively provide a plausible causal account of the quantitative results.

### 5.4. Limitations

Several limitations should be acknowledged. First, the quasi-experimental design, while stronger than a single-group pre-post comparison, does not eliminate all threats to internal validity. Although the two instructors followed a common syllabus and shared an item bank, instructor effects cannot be entirely ruled out. Second, the sample was drawn from a single institution, which limits generalisability. Third, the ERA is a newly developed instrument; while its inter-rater reliability was satisfactory and its construct validity is supported by the alignment of results with the TUCE, further validation studies are needed. Fourth, the study did not include a

delayed post-test, so the durability of CBL effects over time remains an open question. Finally, the qualitative component involved only CBL students; interviews with control-group students would have allowed a fuller comparison of perceived learning experiences.

### 5.5. Implications for Practice

Despite these limitations, the findings carry several implications for economics educators. First, CBL need not replace the lecture; in this study, only one of two weekly contact hours was devoted to cases, and theoretical content coverage was maintained. This "hybrid" model may be more feasible – and less threatening to departmental norms – than a wholesale conversion to case-based instruction. Second, the structured three-phase session format (opening-analysis-synthesis) provides a replicable template that can be adapted to different course levels and topics. Third, the finding that CBL effects are concentrated at higher cognitive levels suggests that programmes should align their assessment strategies accordingly: if the curriculum aims to develop analytical and evaluative reasoning, then assessments must include tasks that elicit those skills, not merely recognition and recall.

## 6. Conclusion

This study investigated the impact of case-based learning on economic reasoning among undergraduate students in a Polish university. The results indicate that a semester-long CBL intervention, integrated into a standard microeconomics course, produced significantly greater gains in economic understanding – particularly at the higher-order levels of application, analysis, and evaluation – compared with a conventional lecture-and-exercise format. Qualitative data suggest that the mechanisms driving these gains include the confrontation with ambiguity, the practice of evidence-based argumentation, and the cognitive benefits of collaborative discussion.

The findings respond to a recognised gap in the economics education literature: the need for controlled studies that employ validated cognitive assessments and that disaggregate outcomes by reasoning level. They also extend the evidence base to a Central European context, where lecture-centred instruction remains the dominant norm. Future research should replicate the design at multiple institutions and across different course levels (introductory, intermediate, advanced), incorporate delayed post-tests to assess the persistence of reasoning gains, and explore the interplay between CBL and other active-learning techniques such as classroom experiments and simulations. Longitudinal tracking of graduates' professional reasoning would further illuminate whether the cognitive benefits observed in the classroom transfer to the workplace. Economics education aspires to produce graduates who can reason about complex, ill-structured problems under conditions of uncertainty. Case-based learning, as the present study demonstrates, is a promising means of turning that aspiration into a measurable pedagogical reality.

## References

Allgood et al., 2015 – *Allgood, S., Walstad, W.B., Siegfried, J.J.* (2015). Research on Teaching Economics to Undergraduates. *Journal of Economic Literature*. 53(2): 285-325. DOI: 10.1257/jel.53.2.285

Anderson, Krathwohl, 2001 – *Anderson, L.W., Krathwohl, D.R.* (Eds.) (2001). A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives. New York: Longman.

Becker, 1997 – *Becker, W.E.* (1997). Teaching Economics to Undergraduates. *Journal of Economic Literature*. 35(3): 1347-1373.

Becker, Watts, 1996 – *Becker, W.E., Watts, M.* (1996). Chalk and Talk: A National Survey on Teaching Undergraduate Economics. *American Economic Review*. 86(2): 448-453.

Becker, Watts, 2001 – *Becker, W.E., Watts, M.* (2001). Teaching Economics at the Start of the 21st Century: Still Chalk-and-Talk. *American Economic Review*. 91(2): 446-451. DOI: 10.1257/aer.91.2.446

Bloom, 1956 – *Bloom, B.S.* (Ed.) (1956). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain. New York: David McKay.

Braun, Clarke, 2006 – *Braun, V., Clarke, V.* (2006). Using Thematic Analysis in Psychology. *Qualitative Research in Psychology*. 3(2): 77-101. DOI: 10.1191/1478088706qp063oa

Buckles, 1998 – *Buckles, S.* (1998). Using Cases as an Effective Active Learning Technique. In: Becker, W., Watts, M. (Eds.). Teaching Economics to Undergraduates: Alternatives to Chalk and Talk. Cheltenham: Edward Elgar: 225-240.

Buckles, Siegfried, 2006 – *Buckles, S., Siegfried, J.J.* (2006). Using Multiple-Choice Questions to Evaluate In-Depth Learning of Economics. *Journal of Economic Education*. 37(1): 48-57. DOI: 10.3200/JECE.37.1.48-57

Carlson, Schodt, 1995 – *Carlson, J.A., Schodt, D.W.* (1995). Beyond the Lecture: Case Teaching and the Learning of Economic Theory. *Journal of Economic Education*. 26(1): 17-28. DOI: 10.1080/00220485.1995.10844853

Carlson, Velenchik, 2006 – *Carlson, J.A., Velenchik, A.D.* (2006). Using the Case Method in the Economics Classroom. In: Becker, W.E., Watts, M., Becker, S.R. (Eds.). Teaching Economics: More Alternatives to Chalk and Talk. Cheltenham: Edward Elgar: 59-74.

Christensen, Hansen, 1987 – *Christensen, C.R., Hansen, A.J.* (1987). Teaching and the Case Method. Boston: Harvard Business School Press.

Emerson, Taylor, 2004 – *Emerson, T.L.N., Taylor, B.A.* (2004). Comparing Student Achievement across Experimental and Lecture-Oriented Sections of a Principles of Microeconomics Course. *Southern Economic Journal*. 70(3): 672-693. DOI: 10.2307/4135338

Freeman et al., 2014 – *Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P.* (2014). Active Learning Increases Student Performance in Science, Engineering, and Mathematics. *Proceedings of the National Academy of Sciences*. 111(23): 8410-8415. DOI: 10.1073/pnas.1319030111

George, Mallery, 2019 – *George, D., Mallery, P.* (2019). IBM SPSS Statistics 26 Step by Step: A Simple Guide and Reference (16th ed.). New York: Routledge.

Goffe, Kauper, 2014 – *Goffe, W.L., Kauper, D.* (2014). A Survey of Principles Instructors: Why Lecture Prevails. *Journal of Economic Education*. 45(4): 360-375. DOI: 10.1080/00220485.2014.946547

Landis, Koch, 1977 – *Landis, J.R., Koch, G.G.* (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 33(1): 159-174. DOI: 10.2307/2529310

Popescu, 2014 – *Popescu, D.* (2014). Teaching Economics – The Case Method. *Annals of the "Constantin Brancusi" University of Targu Jiu, Economy Series*. 6: 274-277.

Ray, 2018 – *Ray, M.* (2018). Teaching Economics Using 'Cases' – Going Beyond the 'Chalk-and-Talk' Method. *International Review of Economics Education*. 27: 1-9. DOI: 10.1016/j.iree.2017.12.001

Siegfried, Walstad, 1998 – *Siegfried, J.J., Walstad, W.B.* (1998). Research on Teaching College Economics. In: Walstad, W.B., Saunders, P. (Eds.). Teaching Undergraduate Economics: A Handbook for Instructors. New York: Irwin/McGraw-Hill: 141-166.

Siegfried et al., 1991 – *Siegfried, J.J., Bartlett, R.L., Hansen, W.L., Kelley, A.C., McCloskey, D.N., Tietenberg, T.H.* (1991). The Status and Prospects of the Economics Major. *Journal of Economic Education*. 22(3): 197-224. DOI: 10.1080/00220485.1991.10844711

Simkins, Maier, 2010 – *Simkins, S., Maier, M.* (Eds.) (2010). Just-in-Time Teaching: Across the Disciplines, Across the Academy. Sterling, VA: Stylus Publishing.

Velenchik, 1995 – *Velenchik, A.D.* (1995). The Case Method as a Strategy for Teaching Policy Analysis to Undergraduates. *Journal of Economic Education*. 26(1): 29-38. DOI: 10.1080/00220485.1995.10844855

Volpe, 2015 – *Volpe, G.* (2015). Case Teaching in Economics: History, Practice and Evidence. *Cogent Economics and Finance*. 3(1): 1120977. DOI: 10.1080/23322039.2015.1120977

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

Walstad et al., 2007 – *Walstad, W.B., Watts, M., Rebeck, K.* (2007). Test of Understanding of College Economics: Examiner's Manual (Fourth Edition). New York: National Council on Economic Education.

Watts, Schaur, 2011 – *Watts, M., Schaur, G.* (2011). Teaching and Assessment Methods in Undergraduate Economics: A Fourth National Quinquennial Survey. *Journal of Economic Education*. 42(3): 294-309. DOI: 10.1080/00220485.2011.581956