



IBonomics Learning Report

Educator-facing analytics for formative assessment, retrieval practice, and targeted revision

June 04, 2026

This report is prepared by the IBonomics team to help teachers and students turn quiz activity into actionable learning evidence. It summarizes where students appear confident, where questions create difficulty, and how those patterns can support targeted teaching, retrieval practice, and revision.

The analysis is designed for educational interpretation rather than student ranking. It treats quiz data as formative feedback: a way to identify concepts, question types, and study routines that may need clearer explanation, more deliberate practice, or better scaffolding.

Prepared by the IBonomics team

Table of Contents

1. Executive Summary	3
1.1 Key Findings	3
1.2 Recommended Actions	3
2. Question Type Overview	4
2.1 Analysis Snapshot	4
3. Data Handling	5
4. Methodology	5
5. Results	6
5.1 Reliable Question Count by Quiz Format	6
5.2 Unit Difficulty Distribution by Main Topic	7
5.3 Specific Topic Difficulty Map	8
5.4 Correct Rate by Question Type	9
5.5 Difficulty Score vs Average Time	10
5.6 Main Topic Difficulty Regression	11
5.7 Hardest Concepts	12
6. Conclusion And Recommendations	13
6.1 For Educators	13
6.2 For Students	13
6.3 Limitations	13

1. Executive Summary

Across 726 reliable IBonomics quiz questions and 44,850 attempts, students answered about 80% correctly on average, suggesting generally secure recall with pockets of substantial difficulty. The toughest items were concentrated in calculation-heavy and trade/protection contexts, and the slowest responses also appeared where formula application and incidence reasoning were required.

1.1 Key Findings

- Calculation items were the weakest question type, with lower correctness than diagram, conceptual, definition, and application items, indicating that procedural fluency is less secure than recall or interpretation.
- The hardest concepts clustered around elasticity and tax/trade mechanics, including PED/PES formula application, VAT incidence, world price change, and required reserve ratio, and these items also took notably longer to answer.
- By main topic, difficulty differences were small and not statistically clear between Introduction to Economics, Microeconomics, Macroeconomics, and Global Economics, so weaknesses appear more concept-specific than topic-wide.

1.2 Recommended Actions

- Use brief retrieval practice on formula selection and substitution for elasticity, tax incidence, and money multiplier questions, then follow with one or two exam-style calculations to reduce processing load.
- Revisit trade protection and supply-side policy items with worked examples that explicitly trace cause-and-effect changes, especially net imports, quotas, world prices, and exchange-rate movements.
- For students, target mixed-topic revision around the identified weak concepts rather than whole units, and check answers by explaining the economic chain in words before moving to the next question.

The report treats quiz results as formative evidence. Its purpose is to identify where content knowledge, task format, fluency, or cognitive load may be limiting performance.

2. Question Type Overview

IBonomics uses several quiz formats that test different learning demands. Unit multiple-choice quizzes emphasize conceptual recall and application and can appear in predefined unit quizzes, custom unit quizzes, or both. Definition quizzes test precise economic vocabulary; math quizzes test procedural fluency; and interactive shift quizzes test diagram interpretation, causal reasoning, and transfer between text and graph.

This distinction matters educationally because a low score can reflect different causes: weak recall, fragile conceptual understanding, calculation errors, graph-reading difficulty, or overload from multi-step reasoning.

2.1 Analysis Snapshot

Metric	Value
Data period	2025-07-17 to 2026-06-02
Reliable analysis sample	726 questions, 44850 question attempts
Reliability threshold	At least 10 cleaned attempts per question

3. Data Handling

The analysis uses question-level quiz response data from the IBonomics database. Only quiz types with stored per-question outcomes are included in this version of the report.

The IB Economics syllabus is organized into broad areas and syllabus units within those areas. In this report, **Unit** refers to those syllabus units, while **Specific Topic** refers to a narrower question-level concept found within a unit. This structure follows the 2022 DP Economics syllabus: [IB Economics SL guide](#) and [IB Economics HL guide](#).

Questions with fewer than 10 cleaned attempts are excluded from report-facing charts, regressions, summaries, and diagnostics. The current audit includes 726 reliable questions and excludes 451 low-attempt questions.

Rows with impossible timing values or incomplete question-level information are removed before analysis. Question times below 1 second and above 300 seconds are treated as invalid so that accidental clicks, unfinished records, and extreme timing artifacts do not distort the report.

Known invalid or confusing question texts are also excluded, thus the sample does not represent all possible questions currently available. In this run, 25 question-attempt rows were removed by that blacklist rule.

User identifiers are excluded from report outputs. The report focuses on aggregate question, topic, and question-type patterns rather than individual student records.

4. Methodology

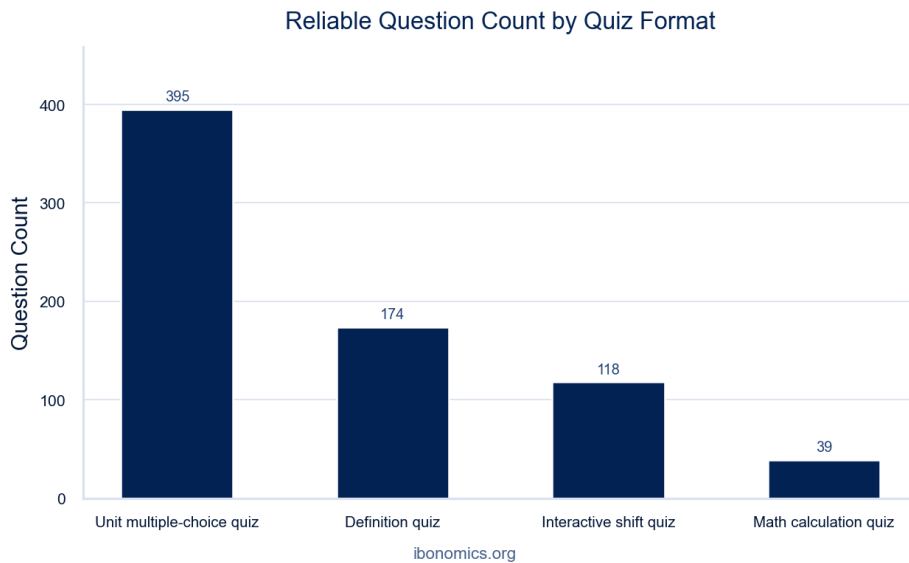
Questions are grouped by question text. Difficulty is calculated as $1 - \text{Correct \%}$, so a higher Diff.Score means lower observed correctness. Time metrics are kept separately because accuracy difficulty and processing time capture different forms of struggle. Combining the two creates four useful question types: easy-fast, easy-slow, hard-fast, and hard-slow. For example, a hard-fast question may indicate a misconception or trap, while an easy-slow question may be calculation-heavy but ultimately understood.

Broad topic and unit labels come from IBonomics curriculum metadata, not from the LLM. Structured LLM outputs with validation add a narrower question taxonomy: Specific Topic identifies the topic within a unit, Concept names the exact idea being tested, Question Type describes the task format, Command Skill describes the action required, and reasoning or misconception fields describe likely cognitive demands. These labels are analytical categories rather than official IB syllabus headings.

The report presents a simple question-level OLS regression that estimates average difficulty-score differences by main topic. The model is descriptive and associative, not causal, and is included to show whether broad syllabus areas alone explain much variation in observed difficulty.

5. Results

5.1 Reliable Question Count by Quiz Format



The bar chart compares the number of reliable unique questions across four IBonomics quiz formats. Unit multiple-choice quizzes have the largest pool, while math calculation quizzes have the smallest.

Key insights

- Unit multiple-choice quizzes dominate the reliable question pool with 395 questions.
- Math calculation quizzes are much more limited at 39 questions, so comparisons here rest on a narrower base.

Action

Use the larger formats for broader practice and the smaller formats for targeted review rather than strong trend comparisons.

Interpretation note: Formats with fewer reliable questions should be interpreted more cautiously because the question base is smaller.

5.2 Unit Difficulty Distribution by Main Topic



Representative High-Difficulty Questions				
Unit	Example Question	Att.	Diff.	Time
Elasticity of Supply	<i>The price of laptops falls from \$1,000 to \$950. If the price elasticity of supply is 0.8 and producers originally supplied 20,000 units, calculate the new quantity supplied.</i>	21	0.952	87.8s
Supply-side Policies	<i>What is the likely economic effect when a major natural disaster destroys factories and infrastructure?</i>	73	0.534	32.5s
Economic Methodology a...	<i>Which concept is associated with Karl Marx's critique of capitalism?</i>	67	0.702	17.5s
Market Failure: Public...	<i>Which of the following is an example of a quasi-public good?</i>	111	0.622	18.6s

The chart compares average difficulty across IB Economics units within each main topic, with point size showing how many reliable questions each unit has. Difficulty varies most within Microeconomics and Global Economics, where some units sit well above or below their topic mean.

Key insights

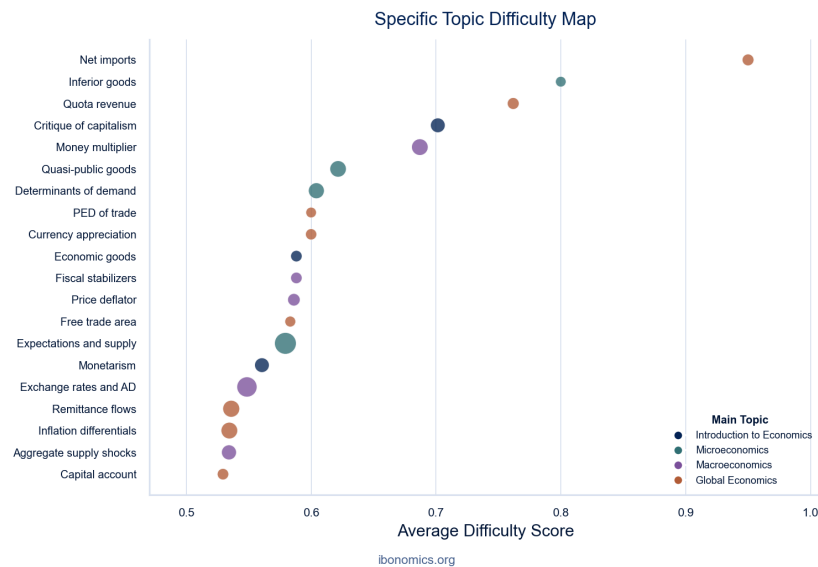
- Exchange Rates is the hardest Global Economics unit at 0.246, well above its main-topic mean of 0.140.
- Within Microeconomics, Elasticity of Supply (0.334) stands out as much harder than the topic mean of 0.205, while several market-failure units are notably lower.

Action

Use the highest-difficulty units as targeted retrieval-practice and review priorities before timed essays or topic tests.

Interpretation note: Units with fewer reliable questions may look more extreme, so treat very small points cautiously.

5.3 Specific Topic Difficulty Map



Representative High-Difficulty Questions				
Specific Topic	Example Question	Att.	Diff.	Time
Inferior goods	<i>In IB Economics, how is an inferior good defined?</i>	10	0.800	8.0s
Critique of capitalism	<i>Which concept is associated with Karl Marx's critique of capitalism?</i>	67	0.702	17.5s
Money multiplier	<i>A commercial bank receives an initial deposit of \$2,000 and the required reserve ratio is 20%. Using the money multiplier formula, calculate by how much the total money supply increases.</i>	45	0.822	57.1s
Quasi-public goods	<i>Which of the following is an example of a quasi-public good?</i>	111	0.622	18.6s

This point map highlights the highest-difficulty specific IB Economics topics, with point size showing how often each topic was attempted. The hardest items are concentrated in Global Economics and Macroeconomics.

Key insights

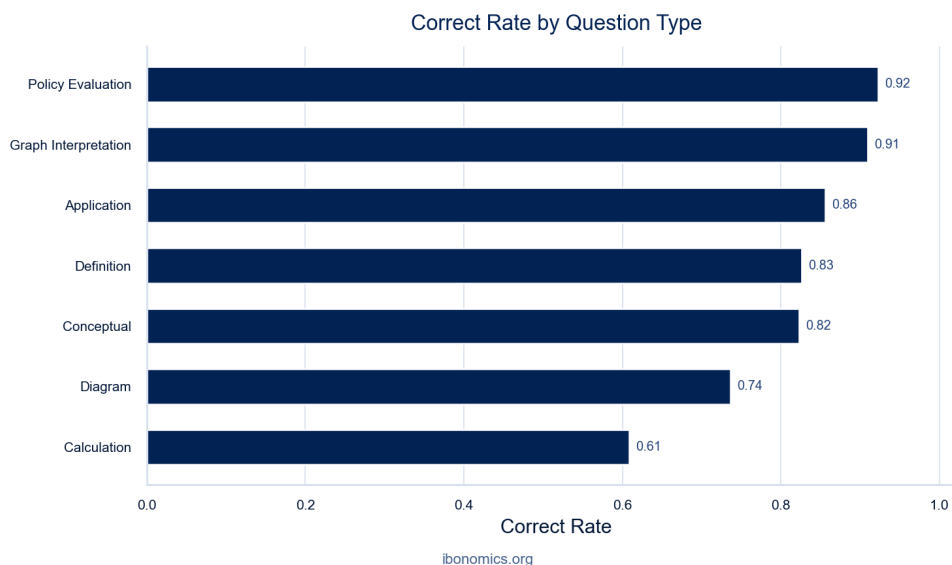
- Net imports is the highest-difficulty topic shown (0.95), though it has only 20 attempts.
- Several high-volume topics remain difficult, especially Expectations and supply (342 attempts, 0.5794) and Exchange rates and AD (257 attempts, 0.5486).

Action

Use this map to prioritise retrieval practice and short review sets on high-volume weak topics before moving to lower-volume outliers.

Interpretation note: Average difficulty is based only on topics with at least 10 cleaned attempts, so very high or low points with small counts should be read cautiously.

5.4 Correct Rate by Question Type



Representative Lower-Correctness Questions				
Question Type	Example Question	Att.	Diff.	Time
Calculation	<i>The price of laptops falls from \$1,000 to \$950. If the price elasticity of supply is 0.8 and producers originally supplied 20,000 units, calculate the new quantity supplied.</i>	21	0.952	87.8s
Diagram	<i>In a single diagram for IB Economics revision courses, show what happens when more EdTech firms enter the market while a new, very effective free online resource reduces students' willingness to pay for paid courses.</i>	136	0.699	44.5s
Conceptual	<i>If a supply curve is flat, what does this indicate about PES?</i>	70	0.700	23.0s
Definition	<i>In IB Economics, how is an inferior good defined?</i>	10	0.800	8.0s

This bar chart compares correct rates across IB Economics question types over the reporting period. Performance is highest on policy evaluation and graph interpretation, and lowest on calculations.

Key insights

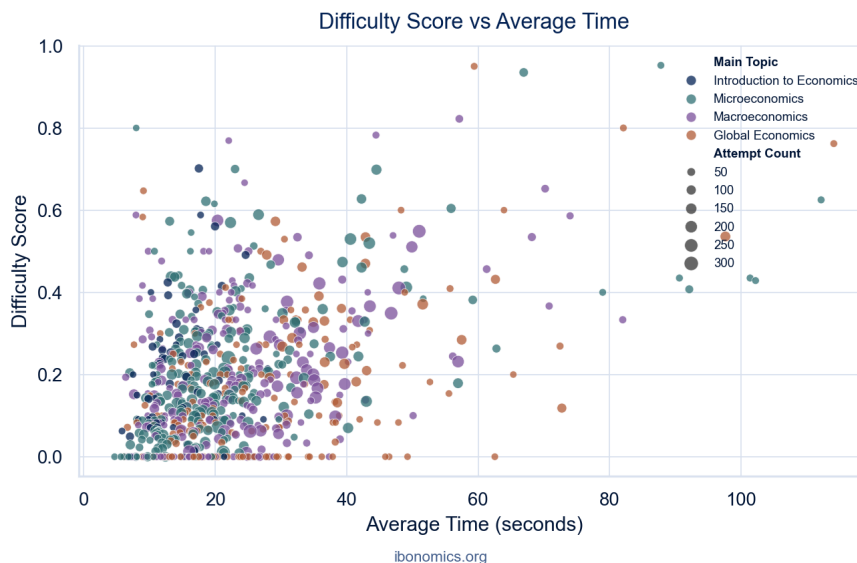
- Policy Evaluation (0.9231) and Graph Interpretation (0.9098) are the strongest areas.
- Calculation is the weakest area at 0.6086, with Diagram also lower at 0.7363.

Action

Use targeted practice sets for calculations and diagrams, then mix them with higher-performing question types in retrieval practice.

Interpretation note: Correct rates show observed performance only and may reflect differences in question difficulty or mix, not skill alone.

5.5 Difficulty Score vs Average Time



Representative Slow-And-Difficult Questions				
Specific Topic	Example Question	Att.	Diff.	Time
Price elasticity of su...	The price of laptops falls from \$1,000 to \$950. If the price elasticity of supply is 0.8 and producers originally supplied 20,000 units, calculate the new quantity supplied.	21	0.952	87.8s
Price elasticity of de...	The price of a pair of brand name headphones falls from \$20 to \$18, and weekly sales rise from 120 to 140 units. Calculate the price elasticity of demand (midpoint method) and state whether total revenue will rise or fall.	24	0.625	112.2s
Import quota	Determine the change in the value of imports at the world price caused by the quota.	20	0.800	82.1s
Indirect tax	The UK applies a 20% VAT on clothing purchases, but no indirect taxes are imposed earlier in the fast fashion production process. If the average person in the UK spends £76.53 per month on clothing, estimate the total VAT paid per person annually on clothing.	77	0.935	66.9s

This scatter plot compares average time against difficulty score for IB Economics questions with at least 10 cleaned attempts, showing where items are both slow and hard versus quick and easier. The spread suggests timing and accuracy do not always move together.

Key insights

- Several trade-protection and elasticity items sit in the slower, harder range, including quota revenue and price elasticity of demand.
- Some of the most difficult items are not the slowest, such as indirect tax and net imports, so accuracy issues can appear even when response time is moderate.

Action

Use upper-right items as worked examples first, then follow with short retrieval practice on the same concepts.

Interpretation note: Average time may reflect reading, calculation, or hesitation, so it should not be treated as a pure measure of conceptual difficulty.

5.6 Main Topic Difficulty Regression

Main Topic	Coef.	Std. Err.	p	95% CI	n	R2
Microeconomics	0.003	0.026	0.892	-0.047 to 0.054	726	0.009
Macroeconomics	0.001	0.026	0.969	-0.049 to 0.051	726	0.009
Global Economics	-0.035	0.026	0.188	-0.086 to 0.017	726	0.009

This regression table compares average difficulty-score differences across main topics relative to Introduction to Economics. All topic coefficients are very close to zero and the model explains little of the variation in difficulty.

Key insights

- Microeconomics and Macroeconomics are essentially indistinguishable from the reference topic in average difficulty.
- Global Economics is slightly lower on average, but the estimate is small and not statistically clear.

Action

Use topic-specific review to target weak subskills, since main topic alone is not a useful predictor of question difficulty.

Interpretation note: These are average differences only; low R² means most difficulty variation comes from factors beyond main topic.

5.7 Hardest Concepts

Concept	Specific Topic	Difficulty	Questions
PED/PES formula application	Price elasticity of supply	0.95	1
World price change	Net imports	0.95	1
VAT incidence	Indirect tax	0.94	1
Required reserve ratio	Money multiplier	0.82	1
Demand falls as income rises	Inferior goods	0.80	1
Import expenditure	Import quota	0.80	1
Reserve ratio	Money multiplier	0.78	1
Expansionary monetary policy	Monetary policy stance	0.77	1
Producer revenue gain	Quota revenue	0.76	1
Alienation	Critique of capitalism	0.70	1

Representative Questions From Hard Concepts

Concept	Example Question	Att.	Diff.	Time
PED/PES formula applic...	<i>The price of laptops falls from \$1,000 to \$950. If the price elasticity of supply is 0.8 and producers originally supplied 20,000 units, calculate the new quantity supplied.</i>	21	0.952	87.8s
VAT incidence	<i>The UK applies a 20% VAT on clothing purchases, but no indirect taxes are imposed earlier in the fast fashion production process. If the average person in the UK spends £76.53 per month on clothing, estimate the total VAT paid per person annually on clothing.</i>	77	0.935	66.9s
Required reserve ratio	<i>A commercial bank receives an initial deposit of \$2,000 and the required reserve ratio is 20%. Using the money multiplier formula, calculate by how much the total money supply increases.</i>	45	0.822	57.1s
Demand falls as income...	<i>In IB Economics, how is an inferior good defined?</i>	10	0.800	8.0s

The table ranks the hardest concepts in the reliable sample, with very low correct rates on several micro, macro, and global economics items. The most difficult entries are formula and incidence topics, plus trade-protection and money-multiplier concepts.

Key insights

- PED/PES formula application and world price change both have about 5% correct rates, indicating very weak performance on these items.
- VAT incidence, reserve ratio, and expansionary monetary policy also score poorly, showing difficulty across both numerical and conceptual topics.

Action

Use these items as a targeted revision checklist: reteach the weakest concepts, then assign short retrieval quizzes on each one.

Interpretation note: Each row is based on a single question, so the table flags likely weak areas rather than stable topic-wide difficulty.

6. Conclusion And Recommendations

The main practical value of the analysis is diagnostic. Reliable question-level patterns can help teachers choose where to reteach, where to add retrieval practice, and where students need more worked examples or format-specific practice.

6.1 For Educators

- Use high-difficulty units as candidates for short retrieval warm-ups and targeted revision lessons.
- Use low-correctness question types to distinguish content gaps from task-format gaps.
- Use slow-and-difficult questions for worked examples, think-aloud modeling, or misconception checks.

6.2 For Students

- Study by specific topic and concept, not only by broad unit.
- Review mistakes by identifying whether the issue was recall, calculation, graph interpretation, or multi-step reasoning.
- Revisit difficult concepts with spaced retrieval rather than repeated immediate retakes only.

6.3 Limitations

The analysis is observational and should be interpreted as formative evidence, not causal evidence. The sample reflects the current IBonomics user population, quiz-selection behavior, repeat practice patterns, and the available question bank. Some units, specific topics, or question types may have small sample sizes even after the reliability threshold, so their estimates are less stable than categories with many questions and attempts.

There is also a risk of sample contamination. Difficulty estimates can be affected by repeated attempts, students sharing answers, abandoned sessions, mistimed attempts, incorrect answer keys and confusing wording. The cleaning rules, timing filters, reliability threshold, and blacklist reduce these risks, but they cannot remove every source of measurement error.

LLM-generated taxonomy labels are validated structurally but are not a substitute for expert curriculum review. Specific Topic, Concept, Question Type, Command Skill, reasoning features, misconception fields, and revision actions should be treated as reproducible analytical labels that help organize the data, not as definitive ground truth. They should be periodically audited for taxonomy drift, inconsistent granularity, or misclassification as new questions and quiz formats are added.