

Commonwealth of Tessaly 2025 (B) — EUR 2036 Notes

CDSBench #260 · Does the instrument's maturity satisfy a maximum maturity of 30 years measured from the credit event date (2025-03-25)?

Commonwealth of Tessaly · Sovereign · Restructuring 2025-03-25 · Rule application · Maturity · Ground truth: deterministically verified

VERIFIED OUTCOME

Yes

Maturity 2036-03-25 vs limit 2055-03-25.

Supporting sections: *DOC-1#S3, DOC-2#S1*

WHAT HAPPENED

All 4 participants reached the verified outcome.

Decided by analyst: Based on the maturityDate provision, the answer is YES.

MODEL COMPARISON

PARTICIPANT	ANSWER	RESULT	CONFIDENCE	EVIDENCE SCORE	PRIMARY FAILURE
Claude	Yes	Correct	75%	67%	—
GPT	Yes	Correct	92%	0%	Evidence selection failure
Gemini	Yes	Correct	85%	0%	Evidence selection failure
Mastermind (reference)	Yes	Correct	75%	67%	—

GPT

GPT: Evidence selection failure

Correct answer but cited evidence does not match the supporting sections (evidence score 0.00).

Model summary: Based on the maturityDate provision, the answer is YES.

Cited: *DOC-1#S1*

GEMINI

Gemini: Evidence selection failure

Correct answer but cited evidence does not match the supporting sections (evidence score 0.00).

Model summary: Based on the maturityDate provision, the answer is YES.

Cited: *DOC-1#S1*

EVIDENCE

Decisive sections of the record

- DOC-1#S3 — EUR 2036 Notes — Terms and Conditions (excerpt), Maturity: The EUR 2036 Notes mature on 2036-03-25. Interest is payable semi-annually in arrear.
- DOC-2#S1 — Determinations Committee notice — Commonwealth of Tessaly, Determination: The Determinations Committee resolved that a Restructuring Credit Event occurred with respect to Commonwealth of Tessaly on 2025-03-25.

METHODOLOGY

A response is correct when its normalised answer equals the verified label. Evidence score is the F1 between cited section IDs and the supporting sections; cited IDs that do not exist count as hallucinated. Abstentions are correct only when the label calls for them. Ground truth is verified by two independent extractors plus an evidence check; no model grades itself.

Full Case Replay with documents and every model's reasoning: <https://cdsbench.com/cases/commonwealth-of-tessaly-2025-b/eur-2036/maturity-limit>