

Halvard Steel AG 2028 (C) — CHF 2064 Notes

CDSBench #289 · Does the instrument's maturity satisfy a maximum maturity of 30 years measured from the credit event date (2028-06-07)?

Halvard Steel AG · Corporate · Restructuring 2028-06-07 · Rule application · Maturity · Ground truth: deterministically verified

VERIFIED OUTCOME

No

Maturity 2064-06-07 vs limit 2058-06-07.

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 NO.

MODEL COMPARISON

PARTICIPANT	ANSWER	RESULT	CONFIDENCE	EVIDENCE SCORE	PRIMARY FAILURE
Claude	No	Correct	75%	67%	—
GPT	No	Correct	92%	0%	Evidence selection failure
Gemini	No	Correct	85%	0%	Hallucinated evidence
Mastermind (reference)	No	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 NO.

Cited: DOC-1#S1

GEMINI

Gemini: Hallucinated evidence

1 cited section id(s) do not exist in the case record.

Model summary: Based on the maturityDate provision, the answer is NO.

Cited: DOC-1#S1, DOC-9#S9 (1 non-existent)

EVIDENCE

Decisive sections of the record

- DOC-1#S3 — CHF 2064 Notes — Terms and Conditions (excerpt), Maturity: The CHF 2064 Notes mature on 2064-06-07. Interest is payable semi-annually in arrear.
- DOC-2#S1 — Determinations Committee notice — Halvard Steel AG, Determination: The Determinations Committee resolved that a Restructuring Credit Event occurred with respect to Halvard Steel AG on 2028-06-07.

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/halvard-steel-ag-2028-c/chf-2064/maturity-limit>