Online Appendix:
"Can ChatGPT Decipher Fedspeak?"

Contents

1 Examples of classified sentences **2**

2 Prompts **3**

2.1 Policy stance classification 3

2.2 Narrative analysis 3

2.2.1 Contractionary U.S. shocks 3

2.2.2 Hindsight Experiment 7

2.2.3 Expansionary U.S. shocks 9

1 Examples of classified sentences

Classification	Sentence
Dovish	The recent softness in inflation is a concern, and the Committee stands ready to act if necessary to ensure that inflation returns to its target over the medium term.
Dovish	The risks to the economic outlook remain tilted to the downside, and the Committee will closely monitor incoming data for any signs of a deterioration in the outlook.
Mostly dovish	The recent easing of financial conditions is welcome, and should help support the economic expansion over the medium term.
Mostly dovish	Incoming data suggest that the economy is performing well overall, but there are some areas of concern, including ongoing trade tensions and a slowing global economy.
Mostly dovish	Given the current state of the economy, the Committee believes that it will be appropriate to maintain the current target range for the federal funds rate for the foreseeable future.
Neutral	The balance of risks to the economic outlook appears roughly balanced at this time.
Neutral	The unemployment rate remained steady at 4.5%
Mostly hawkish	With the unemployment rate at historic lows and inflation near our target, we believe that some further tightening of monetary policy may be warranted in the coming months.
Mostly hawkish	The Committee is aware that financial imbalances could pose risks to the outlook, and we will be vigilant in monitoring these risks going forward.
Mostly hawkish	While the recent slowdown in economic growth is a concern, we believe that it is largely due to transitory factors and expect growth to pick up in the coming quarters.
Hawkish	Given the current state of the economy, the Committee believes that it will be appropriate to continue raising the target range for the federal funds rate at a gradual pace.
Hawkish	We believe that the risks to the economic outlook are skewed to the upside, and we will act preemptively if necessary to ensure that inflation remains under control.
Hawkish	While we acknowledge that some risks to the outlook remain, we believe that the current stance of monetary policy is too accommodative and that further tightening may be needed to keep inflation in check.

Notes: Examples of classifications provided to the human reviewers.

2 Prompts

2.1 Policy stance classification

Policy stance classification Prompt

Classify the following sentence into one of the 5 categories (hawkish, mostly hawkish, neutral, mostly dovish, dovish): [insert sentence]

2.2 Narrative analysis

The LLM pipeline runs in two stages. The first stage is a full-text screen in which five models (GPT-4o, GPT-4o mini, gpt-oss-120b, DeepSeek V4 Pro, and GLM-5.2) independently classify every meeting record as a shock or not. Meetings for which a majority (three out of five) of the models accepts as a shock are clustered into candidate episodes. These episodes enter the second stage, where an adjudicating model (Claude Fable 5) makes a final call. We reproduce the prompts used in these two stages verbatim.

2.2.1 Contractionary U.S. shocks

First-stage screening prompt

You are an expert monetary economist analyzing FOMC meeting transcripts to identify monetary policy shocks following the Romer & Romer (1989) methodology.

A CONTRACTIONARY MONETARY POLICY SHOCK occurs when ALL of the following are true:

1. POLICY ACTION: The Fed tightened monetary policy (raised interest rates, reduced money growth, or increased reserve requirements)
2. INFLATION MOTIVATION: The primary motivation was concern about inflation being TOO HIGH (not just preventing it from rising)
3. ECONOMY AT POTENTIAL: Policymakers believed the economy was roughly at potential/normal

output (not overheating or in recession)

4. NOT BOOM RESPONSE: The action was NOT primarily a response to excessive economic growth or output exceeding potential
5. INDEPENDENCE: The decision reflects changed preferences about acceptable inflation, NOT just responding to new forecasts of real activity

Read the transcript carefully and answer:

CLASSIFICATION: SHOCK or NOT SHOCK

CONFIDENCE: High / Medium / Low

EVIDENCE FOR EACH CRITERION:

1. Policy Action: [Yes/No – quote evidence]
2. Inflation Motivation: [Yes/No – quote evidence]
3. Economy at Potential: [Yes/No – quote evidence]
4. Not Boom Response: [Yes/No – quote evidence]
5. Independence: [Yes/No – quote evidence]

REASONING: [2–3 sentences explaining your classification]

Be concise but thorough. Use actual quotes from the transcript.

Second-stage abjudication prompt

You are adjudicating candidate monetary-policy shock EPISODES from Federal Open Market Committee meeting records. Your job is to reduce false positives from ordinary tightening cycles.

Definition. A contractionary monetary policy shock episode requires all three of:

1. A concrete contractionary policy action (a discount-rate increase, a firmer directive, tighter reserve

availability, or an equivalent concrete step), not hawkish discussion alone.

2. A central stated motive of reducing inflation from its prevailing level, with the Committee accepting losses in output and employment as the price.

3. The action responds to the prevailing rate of inflation itself, not to current or prospective developments in real activity. An action whose central objective is restraining real activity, cooling a boom, or heading off future inflation that has not yet arrived is excluded. Context language about strong demand, capacity pressure, or overheating does NOT by itself disqualify an action aimed at lowering current inflation: most genuine anti-inflation episodes occur in booms, because booms are when inflation is built. The test is what the action responds to, not what the economy looked like. The action must also not be centrally motivated by balance-of-payments, gold, exchange-market, Treasury-financing, or fiscal concerns, and must not be the mere continuation of a tightening path already underway.

Evidence rules.

- Base every judgment ONLY on the packet evidence: the meeting list and the transcript snippets. The vote counts are a triage signal, not evidence.
- Do NOT use outside knowledge of monetary history, published shock chronologies, or anything that happened after the meeting dates in the packet. If the packet evidence is insufficient, say so and use REJECT_INSUFFICIENT_EVIDENCE or SENSITIVITY_ONLY.
- Every "yes" assessment requires an exact quote with its PDF page reference from the packet.

First try to reject or merge. Before you include a shock episode, actively test these alternatives:

1. NO_TIGHTENING: Did the Committee maintain policy, pause, or only discuss future action?

2. CONTINUATION: Is this just another meeting in a tightening path already underway when the cluster window opens?

3. DEMAND_RESPONSE: Does the record show the action was taken IN RESPONSE TO real activity: to restrain a boom, cool demand growth, or prevent future inflation that had not yet arrived? Strong-demand or overheating language describing the economic context is NOT by itself disqualifying. Apply this veto only when reducing the prevailing inflation rate, at accepted cost to output and employment, is NOT the central stated objective of the action.
4. FORECAST_RESPONSE: Is the motive only future inflation risk, staff forecasts, or incoming data, with current inflation not itself judged too high?
5. OTHER_MOTIVE: Are balance-of-payments, gold, exchange-market, Treasury-financing, fiscal, or market-orderliness concerns central enough to weaken the anti-inflation interpretation?
6. DUPLICATE_EPISODE: Should this cluster be merged with an adjacent candidate cluster that is part of the same policy episode?

Allowed decisions. INCLUDE_SHOCK_EPISODE, MERGE_WITH_ADJACENT_CLUSTER, SENSITIVITY_ONLY, REJECT_NO_TIGHTENING, REJECT_ORDINARY_TIGHTENING, REJECT_INSUFFICIENT_EVIDENCE.

Required output JSON.

```
{
  "cluster_id": "",
  "decision": "",
  "canonical_date": "",
  "included_files": [],
  "merged_with": "",
  "criterion_receipts": {
    "policy_action": {"assessment": "yes/partial/no/unclear", "quotes": [{"file": "", "page": 0, "quote":
      ""}]},
    "inflation_motivation": {"assessment": "yes/partial/no/unclear", "quotes": [{"file": "", "page": 0, "quote
      ": ""}]},
```

```
    "accepted_output_costs": {"assessment": "yes/partial/no/unclear", "quotes": [{" file ": "", "page": 0, "
     quote": ""}]},
    "responds_to_inflation_not_real_activity": {"assessment": "yes/partial/no/unclear", "quotes": [{" file ": "",
      "page": 0, "quote": ""}]},
    "not_continuation_or_duplicate": {"assessment": "yes/partial/no/unclear", "quotes": [{" file ": "", "page":
     0, "quote": ""}]}
  },
  "vetoes": [],
  "reasoning": ""
}
```

If the evidence is mixed, prefer SENSITIVITY_ONLY over INCLUDE_SHOCK_EPISODE. If the best interpretation is continuation of an adjacent cluster, use MERGE_WITH_ADJACENT_CLUSTER rather than INCLUDE_SHOCK_EPISODE.

2.2.2 Hindsight Experiment

The hindsight experiment re-adjudicates twelve episodes using the same stage-two system and task prompts shown above, modified so that the adjudicating model is allowed to see and use data on what happened *after* each meeting window. Relative to the main results, we make three edits and leave everything else (the shock definition, the six rejection tests, and the required output JSON) unchanged:

1. The first evidence rule of the second-stage prompt is amended to license an outcome block; and
2. A block reporting subsequent macroeconomic outcomes is appended to (or prepended to) the packet.

This is implemented by replacing the first bullet of the second-stage prompt's **Evidence rules** by the two bullets below; the remaining evidence rules are unchanged.

Amended evidence rule

- In addition to the meeting documents, this packet contains a summary of SUBSEQUENT MACROECONOMIC OUTCOMES observed after the cluster window. You may take this information into

account in your judgment.

- Base every other judgment on the packet evidence: the meeting list and the transcript snippets. The vote counts are a triage signal, not evidence.

The block below is placed either after the meeting records (variant A) or before them (variant B). Bracketed fields are filled with the relevant figures for each episode.

Outcome block appended to the packet

SUBSEQUENT MACROECONOMIC OUTCOMES (observed after the cluster window)

Over the 24 months following [window–end month YYYY]:

– CPI inflation (12–month rate): [x]% at the window's end, [y]% after 12 months, [z]% after 24 months

– Industrial production, cumulative change over 24 months: [w]%

– Unemployment rate: [u]% at the window's end, peak of [p]% within 24 months

– NBER recession beginning or ongoing within 24 months: [yes/no]

The outcome block is filled under one of three conditions:

- **True future.** The episode's actual 24-month aftermath.
- **Fake future.** A severity-matched aftermath borrowed from a different episode under a fixed rotation across the twelve episodes, so that a dramatic aftermath lands on a quiet episode and vice versa.
- **Counterfactual future.** A sign-inverted aftermath: the inflation path is mirrored around its starting value, the industrial-production and unemployment directions are flipped, and any recession is removed.

Each of the three conditions is run in both presentation variants (outcome block after, then before, the meeting records), yielding 12 episodes $\times$ 3 conditions $\times$ 2 variants $=$ 72 adjudications. The twelve episodes comprise the six contested R&R shocks, the two consensus acceptances (December 1968 and October 1979) as

stability controls, and four clear non-shocks as specificity controls. The main-text table reports, for each episode and condition, the share of the two presentation variants that vote to include the episode as a shock.

2.2.3 Expansionary U.S. shocks

The expansionary leg applies the same two-stage design symmetrically to monetary *easing*.

First-stage screening prompt

You are an expert monetary economist analyzing FOMC meeting transcripts to identify monetary policy shocks following the Romer & Romer (1989) methodology, applied symmetrically to EXPANSIONARY policy.

An EXPANSIONARY MONETARY POLICY SHOCK occurs when ALL of the following are true:

1. POLICY ACTION: The Fed eased monetary policy (lowered interest rates, increased money growth, or reduced reserve requirements)
2. UNEMPLOYMENT MOTIVATION: The primary motivation was concern about unemployment being TOO HIGH (not just preventing it from rising)
3. ECONOMY AT POTENTIAL: Policymakers believed the economy was roughly at potential/normal output (not in recession or overheating)
4. NOT SLUMP RESPONSE: The action was NOT primarily a response to weak economic activity, recession, or output below potential
5. INDEPENDENCE: The decision reflects changed preferences about acceptable unemployment or the unemployment-inflation trade-off, NOT just responding to new forecasts of real activity

Read the transcript carefully and answer:

CLASSIFICATION: SHOCK or NOT SHOCK

CONFIDENCE: High / Medium / Low

EVIDENCE FOR EACH CRITERION:

1. Policy Action: [Yes/No – quote evidence]
2. Unemployment Motivation: [Yes/No – quote evidence]
3. Economy at Potential: [Yes/No – quote evidence]
4. Not Slump Response: [Yes/No – quote evidence]
5. Independence: [Yes/No – quote evidence]

REASONING: [2–3 sentences explaining your classification]

Be concise but thorough. Use actual quotes from the transcript.

Second-stage adjudication prompt

Decide whether this cluster is a genuine EXPANSIONARY Romer-Romer-style shock: a deliberate easing to lower unemployment with the economy roughly at potential, accepting inflation consequences, reflecting changed preferences rather than forecasts. The default answer is no. REJECT if the easing is: a response to recession, slump, or output below potential; the unwinding of a prior disinflation; a financial-stress or crisis response (e.g., LTCM, ZLB-era stimulus with large slack); or ordinary forecast-driven accommodation. MERGE with an adjacent cluster if it is the same policy impulse. SENSITIVITY if evidence is mixed but historically plausible. INCLUDE only if all criteria are clearly met.

Respond with JSON:

```
{"cluster": "<cluster_id>", "decision": "reject|merge|sensitivity|include",
 "merge_target": null, "reasoning": "..."}
```