Strategic Compliance in OPEC

A Capacity-Constrained Cournot Benchmark and Explainable Multi-Agent Framework

Ali Masomzadeh

with contributions from Vandit Rikhi and Jonathan Kenigson

Prepared for consideration by *The AI Journal* Technical manuscript draft – August 2026

Article purpose

This manuscript reconstructs a rough Wolfram Language project as a transparent technical feature. It corrects the country sample, separates descriptive evidence from strategic inference, imposes production-capacity constraints, derives one-shot and repeated-game incentives, and specifies an explainable multi-agent architecture suitable for later reinforcement-learning experiments. The numerical results are illustrative until the annual target series, marginal-cost assumptions, and capacity estimates are independently audited against primary sources.

Abstract

The Organization of the Petroleum Exporting Countries presents a canonical coordination problem: collective production restraint can raise the market price, while an individual producer may gain by expanding output when others remain restrained. This article asks what Nash equilibrium can and cannot explain about production relative to announced targets. It rebuilds an initial Wolfram Language analysis into a capacity-constrained partial-equilibrium model and an explainable multi-agent research framework. The empirical component uses a reconstructed annual panel for nine countries from 2018 through 2023. Because the source target series is not fully traceable to a single official monthly dataset, the observations are treated as target proxies rather than definitive legal quotas. The descriptive evidence shows substantial heterogeneity: some countries remain close to target, some are persistently above it, and others are below target or highly volatile. Those ratios do not by themselves identify strategic intent.

The theoretical component derives the exact capacity-constrained Cournot best response, tests unilateral deviations from the 2023 target profile, and evaluates a grim-trigger repeated-game mechanism. Under an illustrative calibration to the 2023 Brent price, global oil demand, provisional marginal-cost proxies, and revealed-output capacity proxies, every modeled country has a positive one-period gain from an upward deviation, but the magnitude and repeated-game sustainability conditions differ sharply. Simple future punishment is insufficient for several producers under the baseline assumptions. Monitoring, expected penalties, compensation, capacity, and heterogeneous fiscal incentives therefore matter as much as the existence of a Nash equilibrium. The article concludes with a reproducible, learning-ready multi-agent architecture in which fixed strategies, best responses, enforcement, demand shocks, and reinforcement learning can be compared without presenting simulation output as a forecast of governmental behavior.

Key findings

- A production-to-target ratio is a measurement, not a moral diagnosis.
- Capacity constraints prevent implausible quantities and meaningless prices.
- In the illustrative 2023 calibration, all nine modeled countries have positive upward-deviation gains, ranging from about \$1.7 million to \$42.6 million per day.
- Required grim-trigger discount factors range from 0.065 to more than one, so no single enforcement mechanism fits every producer.
- A defensible multi-agent model must expose states, actions, rewards, information assumptions, random seeds, and counterfactuals; fixed rules must not be disguised as autonomous learning.

Keywords: OPEC; production targets; Cournot competition; Nash equilibrium; repeated games; monitoring; multi-agent systems; reinforcement learning; explainable AI.

1 The coordination problem and the limits of a simple answer

When a group of oil exporters reduces aggregate production, the resulting scarcity can support a higher price. Yet each producer receives only part of the collective price benefit while bearing the full opportunity cost of its own withheld barrels. If the other producers stay near their targets, a unilateral increase can be privately attractive. This wedge between group and individual incentives is the starting point of the original project (Masomzadeh 2026) and of the present reconstruction.

The natural language of this problem is non-cooperative game theory. A Nash equilibrium is a profile of actions from which no player can profitably deviate while the other players' actions remain fixed (Nash 1950, 1951). Cournot's quantity game predates Nash's general theorem but embodies precisely that best-response logic: firms choose quantities, the market price depends on total quantity, and each firm internalizes the effect of its own output on price while treating rivals' output as fixed (Cournot 1897). A production agreement can therefore be studied by comparing an announced target profile with each country's profit-maximizing response to the others' targets.

That comparison is useful, but it does not settle whether OPEC is a cartel, whether any particular country intentionally violated an agreement, or whether production changes caused observed movements in Brent prices. The empirical literature has repeatedly found that OPEC conduct changes across periods and countries, and that standard behavioral tests can have little power to distinguish collusive, competitive, dominant-firm, and bureaucratic-production hypotheses (Brémond et al. 2012; Griffin 1985; Gülen 1996; Smith 2005). More structural work emphasizes heterogeneous costs, bargaining, capacity, competitive fringe supply, and strategic regime changes (Behar and Ritz 2017; Golombek et al. 2018; Griffin and Xiong 1997; Okullo and Reynès 2016). The correct question is therefore narrower:

> *Given a transparent set of demand, cost, capacity, and information assumptions, what incentives does*

a Nash-Cournot benchmark assign to each producer, and what additional mechanisms would be required to support the target profile over time?

This article makes four contributions to that question. First, it reviews the supplied country panel, corrects its membership frame, and documents its evidentiary limits. Second, it replaces the unconstrained quantity model with a capacity-constrained residual-demand game. Third, it separates a one-shot deviation calculation from a repeated-game enforcement calculation. Fourth, it turns the draft's fixed strategy simulation into an explicit multi-agent architecture that can support either interpretable rules or reinforcement learning, with distinct validation requirements for each.

2 Data reconstruction and measurement discipline

2.1 What the source panel contains

The supplied notebook reports annual target and actual-production values for 2017–2023, a list of rough marginal production costs, annual Brent prices, and Wolfram Language code for descriptive statistics, Cournot equilibria, payoff matrices, clustering, and repeated-strategy simulations (Masomzadeh 2026). The panel was a valuable starting point, but it was not yet a reproducible research dataset. Several annual target values were attributed collectively to OPEC meeting decisions rather than to page-level source records; the cost figures were described as approximate; and the country list mixed members, exempt producers, and a non-member.

The reconstruction therefore preserves the notebook's numerical inputs while weakening the claims attached to them. In this article, the annual target series is called a *reconstructed target proxy*. It is not represented as a definitive official quota ledger. The production series is likewise used as an annual descriptive series, not as a legal finding of conformity or non-conformity.

2.2 Country-frame correction

The source list included Cameroon, which is not an OPEC member, and omitted Angola, which remained a member through 2023 before withdrawing effective 1 January 2024. It also assigned Congo a 2017 observation even though Congo joined OPEC in 2018; Equatorial Guinea joined in 2017 (Organization of the Petroleum Exporting Countries 2026). The revised analysis consequently:

- removes Cameroon;
- begins the common descriptive window in 2018;
- retains nine countries for which the source provides both targets and production: Saudi Arabia, Iraq, Kuwait, the United Arab Emirates, Algeria, Nigeria, Gabon, Congo, and Equatorial Guinea; and
- does not insert Angola ex post, because doing so would require a separately sourced and reconciled series rather than a silent repair.

Iran, Libya, and Venezuela are outside the modeled target panel because the source draft treated them as exempt or otherwise unsuitable for a direct target comparison. The resulting sample is therefore neither all OPEC nor all OPEC+. Every organization-wide interpretation should be qualified accordingly.

Table 1. Descriptive production-to-target statistics, 2018–2023. Ratios above one indicate production above the reconstructed target proxy; they do not by themselves establish deliberate non-compliance.

Country	Mean (SD)	Range	Years > 1
Saudi Arabia	1.026 (0.051)	0.949–1.087	5/6
Iraq	1.053 (0.060)	1.000–1.160	6/6
Kuwait	1.021 (0.066)	0.963–1.150	3/6
UAE	0.999 (0.030)	0.963–1.036	3/6
Algeria	0.958 (0.028)	0.915–0.996	0/6
Nigeria	0.940 (0.121)	0.770–1.095	2/6
Gabon	1.051 (0.029)	0.994–1.079	5/6
Congo	1.078 (0.124)	0.946–1.236	4/6
Eq. Guinea	0.924 (0.078)	0.857–1.049	1/6

2.3 A neutral production-to-target ratio

For country i and year t, define

$$r_{it} = \frac{y_{it}}{\tau_{it}}, \tag{1}$$

where y_{it} is reported actual production and τ_{it} is the reconstructed target proxy. Then $r_{it} > 1$ means production above the proxy, $r_{it} < 1$ means production below it, and $r_{it} = 1$ means equality after annual aggregation and rounding.

The ratio deliberately avoids labels such as "cheating," "defection," or "over-compliance" in the descriptive section. Those labels require more evidence than the ratio contains. A country can fall below target because mature fields, outages, sanctions, conflict, maintenance, or investment constraints prevent it from producing more. A country can appear above target because of revisions, baseline changes, monthly timing, condensate definitions, or aggregation. Strategy is one possible explanation, not an observation.

Table 1 and figures 1a and 1b reveal three patterns. First, the United Arab Emirates is close to the reconstructed target on average, with a mean ratio of 0.999 and a standard deviation of 0.030. Second, Iraq and Gabon are above target in nearly every year of the common window, although the economic significance of a given ratio depends on the country's scale. Third, Nigeria and Congo are much more volatile than the remaining countries. These are descriptive regularities, not estimates of structural parameters.

The small annual sample is not suitable for the causal price regression attempted in the initial notebook. Six common-year observations cannot separate OPEC output from global demand, non-OPEC production, inventories, financial conditions, sanctions, wars, and the COVID-19 shock. Oil-market identification itself depends sensitively on demand and supply elasticities and inventory behavior (Caldara et al. 2019; Kilian and Murphy 2014). The revised article therefore does not claim that its annual panel identifies a causal effect on Brent.

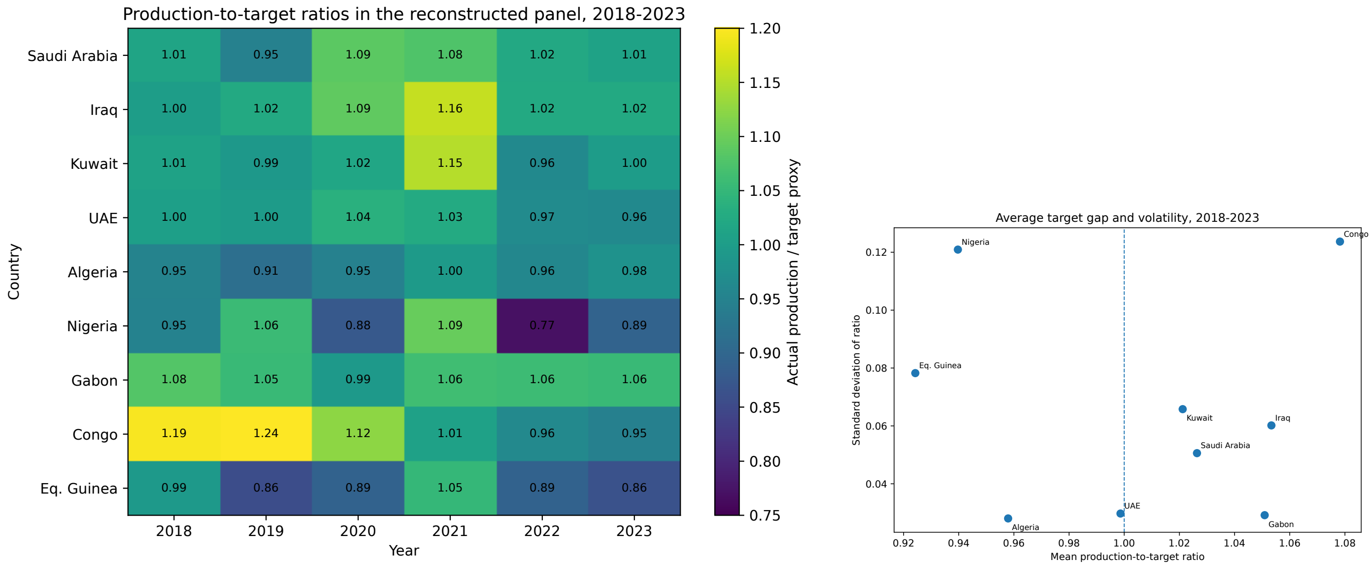
3 A capacity-constrained residual-demand game

3.1 Players, actions, demand, and payoffs

Let $N = \{1, \ldots, n\}$ denote the modeled producers. Country i chooses production

$$q_i \in [0, K_i], \tag{2}$$

where $K_i > 0$ is a short-run production-capacity bound. Total modeled output is $Q = \sum_{j=1}^{n} q_j$. Other oil supply is treated as fixed background supply over the local experiment, so the



(a) Annual production-to-target ratios. The color scale and cell labels report r_{it} from equation (1); they do not infer intent.

(b) Average target gap and year-to-year variability.

Figure 1. Descriptive patterns in the reconstructed 2018–2023 panel. Distance from one measures the direction of the average target gap; vertical position in panel (b) measures volatility.

modeled group faces the residual inverse demand

$$P(Q) = [\alpha - \beta Q]_+, \qquad \alpha > 0,\ \beta > 0. \tag{3}$$

The positive-part operator prevents a negative price outside the local calibration range. All reported scenarios remain in the positive-price region.

Country i has constant short-run marginal cost c_i and one-period profit

$$\pi_i(q_i, Q_{-i}) = \big(\alpha - \beta(q_i + Q_{-i}) - c_i\big) q_i, \qquad Q_{-i} = \sum_{j \neq i} q_j. \tag{4}$$

When q_i is in mb/d and price and cost are in USD per barrel, π_i is in million USD per day. This profit measure is operating profit before fixed costs, taxes, transport, quality differentials, fiscal transfers, and intertemporal scarcity rents. It should not be confused with government fiscal break-even revenue.

Proposition 1 (Capacity-constrained best response)**.** *For fixed rival output Q_{-i}, the profit function in equation (4) is strictly concave in q_i. Country i has the unique best response*

$$\mathrm{BR}_i(Q_{-i}) = \min\left\{K_i, \max\left\{0, \frac{\alpha - c_i - \beta Q_{-i}}{2\beta}\right\}\right\}. \tag{5}$$

Proof. Expanding equation (4) gives

$$\pi_i = (\alpha - c_i - \beta Q_{-i}) q_i - \beta q_i^2.$$

Its second derivative is $-2\beta < 0$. The unconstrained stationary point is

$$q_i^\circ = \frac{\alpha - c_i - \beta Q_{-i}}{2\beta}.$$

Strict concavity and projection of q_i° onto the compact feasible interval $[0, K_i]$ yield equation (5). □

A capacity-constrained Cournot-Nash equilibrium is a vector $q^* \in \prod_i [0, K_i]$ satisfying

$$q_i^* = \mathrm{BR}_i\left(\sum_{j \neq i} q_j^*\right) \quad \text{for every } i. \tag{6}$$

If no capacity or non-negativity constraint binds, the familiar asymmetric interior solution is

$$Q^* = \frac{n\alpha - \sum_{j=1}^n c_j}{(n+1)\beta}, \tag{7}$$

$$q_i^* = \frac{\alpha + \sum_{j=1}^n c_j - (n+1)c_i}{(n+1)\beta}, \tag{8}$$

$$P^* = \frac{\alpha + \sum_{j=1}^n c_j}{n+1}. \tag{9}$$

Those formulas are invalid when they assign negative output or output above physical capacity. The original notebook's most extreme quantities arose precisely because it applied an unconstrained formula to producers of radically different scale.

Corollary 1 (All-capacity equilibrium condition)**.** *Within the positive-price region, the vector $q^* = K = (K_1, \ldots, K_n)$ is a Nash equilibrium if and only if*

$$\alpha - c_i - \beta \sum_{j \neq i} K_j - 2\beta K_i \geq 0 \quad \textit{for every } i. \tag{10}$$

The left side of equation (10) is country i's marginal profit at its capacity when every rival is also at capacity. If it is non-negative, the country would expand further if expansion were feasible; its constrained best response is therefore K_i.

3.2 Testing a target profile rather than inventing a defect quantity

Let $t = (t_1, \ldots, t_n)$ denote a target profile and $T_{-i} = \sum_{j \neq i} t_j$. The target profile is a one-shot Nash equilibrium exactly when

$$t_i = \mathrm{BR}_i(T_{-i}) \quad \text{for every } i. \tag{11}$$

The optimal unilateral deviation is not an unconditional Cournot quantity calculated under a different rival-output profile. It is the conditional best response

$$q_i^D = \mathrm{BR}_i(T_{-i}), \tag{12}$$

and the one-shot deviation gain is

$$\Delta_i = \pi_i(q_i^D, T_{-i}) - \pi_i(t_i, T_{-i}). \tag{13}$$

A positive Δ_i rejects the target profile as a one-shot Nash equilibrium for that country under the maintained assumptions. It does not show that the country actually deviated, knew the parameters, maximized the modeled objective, or could sell the extra quantity without contractual or logistical frictions.

4 Transparent calibration

4.1 Local inverse-demand slope

The model is calibrated as a local comparative-static benchmark around 2023, not as a structural forecast. The U.S. Energy Information Administration reports a 2023 annual average Brent-Europe spot price of $82.49 per barrel (U.S. Energy Information Administration 2026). OPEC's 59th Annual Statistical Bulletin reports average 2023 world oil demand of 102.2 mb/d (Organization of the Petroleum Exporting Countries 2024b). With an assumed absolute short-run demand elasticity $|\eta|$, the local inverse-demand slope is

$$\beta = \frac{P_0}{|\eta| D_0}. \tag{14}$$

This follows from $|\eta| = \left|\frac{dD}{dP}\frac{P_0}{D_0}\right|$ and $\left|\frac{dP}{dD}\right| = \frac{P_0}{|\eta| D_0}$. Holding background supply fixed, a change in modeled-group output shifts the market-clearing quantity along this local demand curve.

The baseline sets $|\eta| = 0.20$ and varies it from 0.10 to 0.30. These values are assumptions, not estimates from the six-year country panel. Oil-demand elasticity is a consequential identification choice, and the literature warns against treating any single value as innocuous (Caldara et al. 2019).

Let $Q_0 = 23.125$ mb/d be the sample's reported 2023 actual output. The intercept is chosen so that $P(Q_0) = P_0$:

$$\alpha = P_0 + \beta Q_0. \tag{15}$$

This is a *residual* demand curve for the modeled nine-country sample. Calibrating the intercept to global demand and then subtracting only sample output, as in the rough draft, double-counts the background supply and generates prices above $300 per barrel at observed quantities. The reconstruction avoids that mismatch.

Table 2. Illustrative baseline calibration. Quantities are million barrels per day (mb/d).

Input	Value (unit)
2023 Brent reference price	82.49 USD/bbl
2023 world oil demand	102.20 mb/d
Absolute short-run demand elasticity	0.20
Modeled-sample actual output	23.125 mb/d
Local inverse-demand slope β	4.036 USD/bbl per mb/d
Residual-demand intercept α	175.816 USD/bbl
Modeled-sample target total	23.260 mb/d
Price at target profile	81.945 USD/bbl
Capacity proxy multiplier	1.05 × historical max.
Modeled-sample capacity total	25.770 mb/d
Price at constrained equilibrium	71.815 USD/bbl

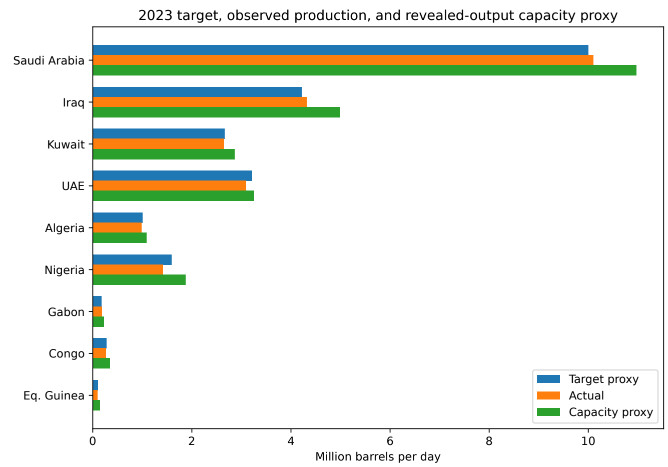


Figure 2. The 2023 target proxy, observed production, and revealed-output capacity proxy. The capacity series is a transparent modeling device, not an engineering estimate.

4.2 Capacity and cost inputs

Engineering capacity estimates were not supplied with the source draft. The baseline therefore uses the transparent revealed-output proxy

$$K_i = m \max_{t \in \{2018, \ldots, 2023\}} y_{it}, \qquad m = 1.05, \tag{16}$$

with sensitivity values $m \in \{1.00, 1.05, 1.10, 1.20\}$. This proxy says only that short-run feasible output is modestly above the highest observed annual average. It is not a reserve estimate, nameplate capacity, sustainable maximum capacity, or estimate of spare capacity.

The marginal-cost values are inherited from the supplied notebook. They are retained to preserve the project's original economic heterogeneity but are treated as rough operating-cost proxies. Before publication as an empirical research claim, each value should be replaced or triangulated with documented field-level or country-level sources and accompanied by uncertainty intervals.

At the reconstructed 2023 target total of 23.260 mb/d, the local model implies a price of $81.945 per barrel, close to the calibration point because target and actual output differ by only 0.135 mb/d. At the baseline capacity total of 25.770 mb/d, the implied price is $71.815 per barrel. These values are internally consistent outputs of the local model; they are not retrospective predictions of Brent because background supply, demand, inventories, quality differentials, and expectations are held fixed.

5 One-shot equilibrium results

Under the baseline calibration, the best response to rivals remaining at their 2023 targets is the capacity proxy for every modeled country. The condition in equation (10) also holds for every country at the all-capacity profile, so the constrained Cournot equilibrium is $q^* = K$. The result should be read structurally: within this local model, the target quantities lie below the countries' short-run profit-maximizing upward choices, and the capacity bounds become decisive.

The sum of the positive one-shot gains is approximately \$118.3 million per day. Iraq has the largest modeled gain at \$42.6 million per day, followed by Saudi Arabia at \$33.2 million and Nigeria at \$14.6 million. The United Arab Emirates has the smallest gain among the large producers, about \$2.1 million per day, because its target proxy is already close to its capacity proxy. The absolute gains for Gabon, Congo, and Equatorial Guinea are small relative to those of Saudi Arabia and Iraq, but their target quantities are also far smaller.

Three interpretations are justified. First, the target profile is not a one-shot Nash equilibrium under this calibration. Second, incentives are heterogeneous: the same verbal label, "produce above target," conceals large differences in quantity, profit, and market impact. Third, the numerical values depend heavily on capacity and demand assumptions. The model does *not* justify the stronger claim that every real producer always wants to exceed every real agreement. A change in capacity, marginal cost, outside supply, demand, fiscal objective, or target allocation can reverse a best response.

The all-capacity equilibrium also illustrates why the word "Cournot" is not itself a guarantee of realism. If the capacity proxy is too high, the equilibrium overstates feasible output; if it is too low, it mechanically suppresses strategic quantity choice. The correct response is not to remove capacity, but to measure it better and report sensitivity.

6 Repeated interaction, monitoring, and enforcement

6.1 A country-specific grim-trigger condition

Suppose the target profile is repeated indefinitely. Let π_i^C be country i's profit at the target profile, π_i^D its profit from the optimal one-period deviation in equation (13), and π_i^P its per-period profit under a credible punishment profile. Let $\delta_i \in [0, 1)$ be its discount factor. Under a grim-trigger strategy, continued target production has present value

$$V_i^C = \frac{\pi_i^C}{1-\delta_i}, \tag{17}$$

whereas a one-time deviation followed by permanent punishment has value

$$V_i^D = \pi_i^D + \frac{\delta_i \pi_i^P}{1-\delta_i}. \tag{18}$$

If $\pi_i^D > \pi_i^P$, the target is incentive-compatible under this strategy when

$$\delta_i \geq \delta_i^* := \frac{\pi_i^D - \pi_i^C}{\pi_i^D - \pi_i^P}. \tag{19}$$

The formula is country-specific. A high deviation gain raises the threshold; a severe punishment lowers it. If $\delta_i^* > 1$, no admissible discount factor can sustain the target with the specified punishment. If $\pi_i^P \geq \pi_i^D$, the punishment is not sufficiently adverse to deter a profitable deviation.

The baseline uses the all-capacity Nash equilibrium as the punishment profile. This is deliberately stark, analogous to a price-war phase in repeated-cartel models (Green and Porter 1984). It is credible within the stage game because it is a constrained Nash equilibrium, but it is not necessarily institutionally realistic or welfare-maximizing.

The results reject a uniform repeated-game story. The United Arab Emirates has $\delta_i^* = 0.065$ because its deviation gain is small and the punishment reduces its profit substantially. Saudi Arabia, Kuwait, and Algeria have thresholds below 0.51. Nigeria requires a much higher 0.869. Congo is nearly at the theoretical boundary, 0.984. Iraq, Gabon, and Equatorial Guinea have thresholds above one: in the baseline, the all-capacity punishment does not make future losses large enough to offset the one-period gain.

For Gabon and Equatorial Guinea, the explanation is especially important. Their capacity quantities are sufficiently above their targets that they can earn more in the punishment equilibrium than at the target profile despite the lower price. A collective price war is therefore not automatically a punishment for every member. Effective coordination may require asymmetric transfers, tailored compensation, access to future target allocations, investment concessions, or other country-specific instruments. This is consistent with the broader repeated-game insight that many outcomes can be supported only when continuation payoffs are designed around heterogeneous incentives (Fudenberg and Maskin 1986).

6.2 Monitoring and an expected marginal penalty

A target cannot be enforced unless output is observed with enough accuracy and enough speed. Real coordination systems therefore distinguish measurement from punishment. OPEC's Joint Ministerial Monitoring Committee reviews production data, emphasizes conformity, and has used compensation plans for outstanding overproduced volumes (Organization of the Petroleum Exporting Countries 2024a). The present model does not reproduce those institutions, but it can show the scale of an expected incentive required to make an upward deviation unattractive.

For a target strictly below capacity, $t_i < K_i$, let $g_i = \rho_i f_i$ denote the expected marginal consequence per excess barrel, where ρ_i is the probability that excess output is detected and acted upon and f_i is the consequence conditional on enforcement. Add the one-sided term

$$-g_i(q_i - t_i)_+ \tag{20}$$

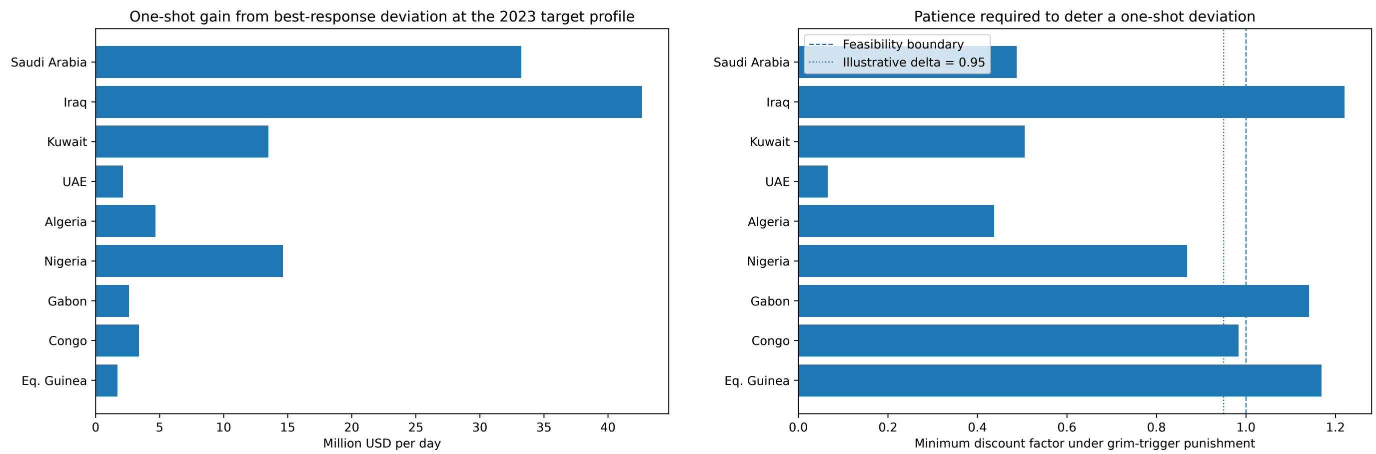
to country i's stage payoff. Because the unpenalized profit is concave, no upward deviation from t_i is profitable if the right derivative at the target is non-positive. The exact threshold is

$$g_i \geq g_i^\dagger := \max\{0, \alpha - c_i - \beta T_{-i} - 2\beta t_i\}. \tag{21}$$

The baseline values in table 3 range from approximately \$38 to \$66 per excess barrel. They are not estimates of an appropriate

Table 3. One-shot deviation and repeated-game results under the baseline calibration. q_i^D is the capacity-constrained best response to the other countries remaining at their 2023 target proxies. δ_i^* is the minimum discount factor under the specified grim-trigger punishment; values above one are economically infeasible. The final column is the minimum expected marginal penalty per excess barrel required to make the target a one-sided optimum against upward deviations.

Country	Target (mb/d)	q_i^D (mb/d)	Gain (million USD/day)	δ_i^*	Penalty (USD/bbl)
Saudi Arabia	10.000	10.973	33.22	0.488	38.09
Iraq	4.220	4.998	42.61	1.220	57.91
Kuwait	2.660	2.868	13.46	0.505	65.71
UAE	3.220	3.255	2.11	0.065	60.45
Algeria	1.000	1.087	4.65	0.437	53.91
Nigeria	1.590	1.869	14.62	0.869	53.53
Gabon	0.180	0.229	2.59	1.140	53.22
Congo	0.280	0.346	3.36	0.984	50.82
Eq. Guinea	0.110	0.146	1.70	1.168	47.50



(a) Modeled one-shot gains from the conditional best response while the other countries remain at their 2023 target proxies.

(b) Minimum discount factors under the specified all-capacity punishment. Values above one are infeasible.

Figure 3. Static and repeated-game incentives under the baseline calibration. Panel (a) measures immediate temptation; panel (b) measures the patience required for the designated punishment to deter it.

legal fine. They are shadow-price outputs of an illustrative model and combine detection, future target adjustments, compensation requirements, reputational consequences, and any other marginal cost of exceeding the target.

Equation (21) also clarifies the value of better information. A more accurate monitoring system can raise ρ_i, reducing the conditional consequence f_i required to reach the same expected marginal deterrent. Conversely, if output is noisy or privately observed, equilibrium enforcement must account for false positives and false negatives, and periods of low price may be statistically indistinguishable from hidden deviations under demand uncertainty (Green and Porter 1984).

7 From fixed strategy scripts to explainable multi-agent AI

7.1 Why the original simulation was not yet adaptive

The supplied notebook defined strategies such as Always Comply, Always Defect, Tit-for-Tat, Grim Trigger, a threshold rule, and a Saudi enforcement rule. It then assigned one rule to each country and iterated the resulting actions. That is a useful *rule-based repeated-game simulation*; it is not yet an adaptive agent-based model in which agents learn, optimize discounted rewards, estimate hidden states, or switch policies from experience.

The distinction matters. A fixed strategy can be informative when its behavioral assumption is explicit. Calling it "agentic" or "AI" does not add validity. Conversely, reinforcement learning can make the model less interpretable and less stable if the state, reward, exploration process, and validation protocol are not transparent. Agent-based computational economics treats macro-patterns as outcomes generated by interacting computational agents (Tesfatsion 2003); prior work has combined evolutionary games and agent-based methods in historical studies of oil-market power (Wood et al. 2016). The present framework retains that bottom-up logic while placing interpretability and falsifiability ahead of simulated complexity.

7.2 A formal learning-ready environment

At period t, let a public environment state be

$$s_t = (P_{t-1},\ G_{t-1},\ M_{t-1},\ z_t,\ e_t)\,, \tag{22}$$

where P_{t-1} is the previous price, $G_{t-1} = \sum_i (q_{i,t-1} - t_{i,t-1})$ is the aggregate target gap, M_{t-1} is the monitored or reported output signal, z_t is a demand-state indicator, and e_t records the enforcement or compensation regime. Country i may also observe private components such as capacity availability, outages, fiscal pressure, or its own cost shock.

An action can be continuous, $a_{it} = q_{it} \in [0, K_{it}]$, or discretized for an interpretable experiment, for example

$$a_{it} \in \{0.95t_{it},\ t_{it},\ 1.025t_{it},\ 1.05t_{it},\ K_{it}\} \cap [0, K_{it}]. \tag{23}$$

The reward should expose every substantive assumption:

$$R_{it} = (P_t - c_{it})q_{it} - g_{it}(q_{it} - t_{it})_+ - \lambda_i(q_{it} - q_{i,t-1})^2 + u_i(s_t, a_t). \tag{24}$$

The first term is operating profit, the second is expected enforcement, the third is an optional adjustment cost, and u_i is an explicitly documented reward component, measured in the same units, for fiscal, political, alliance, or capacity-investment objectives. Nothing should be hidden in an unnamed “agent preference” coefficient.

The market transition is

$$P_t = \left[\alpha(z_t) - \beta(z_t)\sum_i q_{it}\right]_+, \tag{25}$$

$$s_{t+1} \sim \mathcal{T}(\cdot \mid s_t, a_t), \tag{26}$$

where $\mathcal{T}$ governs demand shocks, monitoring error, capacity changes, and enforcement responses. The deterministic model is recovered when $\mathcal{T}$ is degenerate and all policies are fixed rules.

7.3 Three policy families

A rigorous simulation should compare, rather than conflate, three kinds of agents:

1. **Transparent rule agents.** Always-target, always-capacity, Tit-for-Tat, and Grim Trigger provide interpretable reference cases. Their value is diagnostic, not predictive.
2. **Myopic optimizing agents.** Each period, country i uses equation (5), modified for the current state, observed rivals, and enforcement term. This connects simulation directly to the static model.
3. **Learning agents.** A tabular baseline can update the action-value function Q_i using

$$Q_i(s_t, a_{it}) \leftarrow (1 - \ell_i)Q_i(s_t, a_{it}) + \ell_i\left[R_{it} + \gamma_i \max_{a'} Q_i(s_{t+1}, a')\right], \tag{27}$$

where ℓ_i is a learning rate and γ_i a reward-discount factor (Sutton and Barto 2018; Watkins and Dayan 1992). In a multi-agent environment, each learner changes the environment faced by the others, so the single-agent convergence theorem does not automatically apply. Stability must be demonstrated empirically rather than assumed.

Minimal reproducible simulation protocol

```
Input: audited targets, production, costs, capacities, demand
      process,
       monitoring model, enforcement rule, policy class,
             horizon H, random seed.

For each experimental condition and seed:
    initialize country states and policy parameters
    for t = 1 to H:
        publish common state and each country's permitted
              private state
        each country selects an action under its documented
              policy
        clear the market and generate monitored output signals
        apply enforcement and compute decomposed rewards
        update learning policies, if present
        log state, action, price, reward components, and policy
              version
    end
    save full trajectory and summary diagnostics
end

Report: distributions across seeds, policy maps,
      counterfactual decompositions,
        calibration error, out-of-sample error, and failure
              cases.
```

7.4 Explainability is part of the model, not a post-processing graphic

An explainable multi-agent study should make every simulated action reconstructible. For each country-period, the log should identify:

- the state variables available to the agent;
- the feasible action set after capacity and institutional constraints;
- the chosen action and, for stochastic policies, its selection probability;
- operating profit, enforcement, adjustment, and non-economic reward components separately;
- the counterfactual reward from remaining at target and from choosing the myopic best response;
- the policy parameters or model checkpoint used; and
- the random seed and software version.

These records support three levels of explanation. A *local explanation* shows why one country chose one action in one state. A *policy explanation* visualizes how actions change across price, enforcement, and target-gap states. A *system explanation* decomposes an aggregate outcome, such as a price decline, into demand shocks, output changes, capacity changes, and enforcement responses. This is more informative than presenting a smooth line chart whose behavior is predetermined by unreported rules.

8 Sensitivity and robustness

The baseline is only one point in a deliberately simple parameter space. Table 4 and figure 4 vary the absolute demand elasticity and the capacity multiplier. Several conclusions survive: the capacity-constrained equilibrium price is below the target-profile price, positive deviation incentives are common, and repeated-game thresholds differ sharply across countries. The magnitudes do not survive unchanged.

When demand is less elastic, the inverse-demand curve is steeper. Additional output causes a larger price reduction, making an all-capacity punishment more severe and lowering the repeated-game discount thresholds. When demand is more elastic, the price penalty from collective expansion is weaker, so patience alone is less effective. Increasing the capacity multiplier has two competing effects: it raises the one-period quantity available to a deviator, but it also makes the punishment equilibrium more destructive to price. In the reported grid, the latter effect lowers the median δ_i^* as capacity increases, even while aggregate one-shot gains rise.

At the baseline $(|\eta|, m) = (0.20, 1.05)$, all nine countries have positive one-shot gains, the median grim-trigger threshold is 0.869, and five countries can be sustained at an illustrative $\delta = 0.95$. Holding $|\eta| = 0.20$, at $m = 1.00$ only eight have positive gains and the median threshold among those deviators

Table 4. Sensitivity to the absolute short-run demand elasticity and the revealed-output capacity multiplier. The number sustainable at $\delta = 0.95$ is calculated only for countries with positive one-shot deviation gains.

$\lvert\eta\rvert$	Capacity multiplier	Equilibrium price (USD/bbl)	Total deviation gain (million USD/day)	Deviators	Median δ^*	Sustainable at $\delta = 0.95$
0.10	1.00	71.04	39.5	8	0.613	7
0.10	1.05	61.14	63.3	9	0.367	9
0.10	1.10	51.24	91.1	9	0.284	9
0.10	1.20	31.43	141.3	9	0.207	9
0.20	1.00	76.77	68.9	8	1.409	3
0.20	1.05	71.81	118.3	9	0.869	5
0.20	1.10	66.86	172.0	9	0.719	8
0.20	1.20	56.96	270.0	9	0.576	9
0.30	1.00	78.67	79.7	8	2.174	2
0.30	1.05	75.37	140.2	9	1.370	3
0.30	1.10	72.07	206.5	9	1.137	3
0.30	1.20	65.47	332.8	9	0.915	6

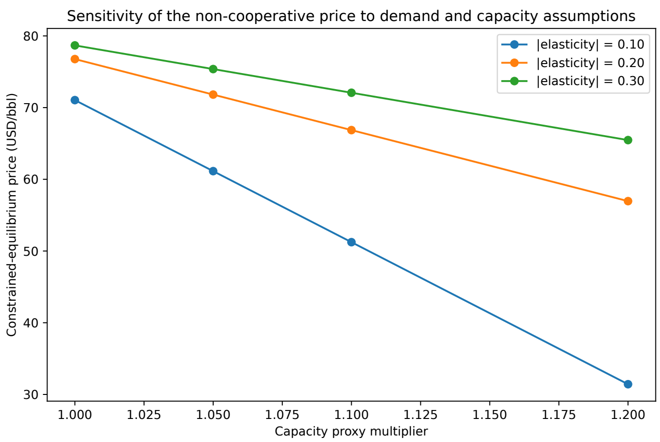


Figure 4. Sensitivity of the constrained-equilibrium price to the local demand elasticity and capacity proxy.

is above one. At $m = 1.10$, eight of nine positive deviators are sustainable at $\delta = 0.95$. These differences are too large to treat the baseline numbers as estimates without independently measured capacities and a richer demand system.

9 What the model establishes – and what it does not

Interpretive boundary

The model identifies incentives conditional on its assumptions. It does not identify intentional misconduct, estimate a causal effect of production on Brent, predict diplomatic decisions, or establish that OPEC behaves as a single stable cartel. Those claims require different data and identification strategies.

The analysis has seven principal limitations.

1. **The target panel is reconstructed.** Annual averages can conceal monthly changes, voluntary cuts, baseline revisions, compensation schedules, and country exemptions. The next empirical version should build a month-by-country panel from page-level official decisions and secondary-source production data.
2. **The sample is incomplete.** Angola is absent because the source draft omitted it; exempt members and non-OPEC OPEC+ participants are outside the game. The residual-demand interpretation is therefore necessary, and organization-wide conclusions are unwarranted.
3. **Capacity is proxied, not measured.** Maximum observed annual production times a multiplier is transparent but crude. Sustainable capacity, temporary surge capacity, field decline, outages, and export infrastructure should be modeled separately.
4. **Costs are rough.** Constant marginal costs omit taxes, royalties, transport, quality, depletion, investment, domestic consumption, and fiscal dependence. A state may maximize an objective different from commercial operating profit.
5. **Outside supply is fixed locally.** U.S. shale, Russia, Canada, Brazil, and other producers respond to price with different lags. Dominant-firm and competitive-fringe models provide richer alternatives (Behar and Ritz 2017; Golombek et al. 2018).
6. **The demand curve is local and linear.** It is a comparative-static approximation around 2023, not a global forecast equation. Inventories, expectations, financial positions, product markets, and geopolitical risk are omitted (Caldara et al. 2019; Kilian and Murphy 2014).
7. **The repeated game is stylized.** Grim Trigger assumes a public, correctly attributed deviation and permanent punishment. Real agreements are renegotiated, monitored imperfectly, and supported by compensation, diplomacy, side payments, future baselines, and asymmetric bargaining.

These limitations are not reasons to abandon the model. They specify the boundary between a useful benchmark and an overclaim. A rigorous benchmark should be simple enough to audit and strong enough to expose which additional facts are needed.

10 A publication-grade empirical roadmap

A stronger second-stage study would replace the reconstructed annual panel with monthly data and preregister the following workflow.

1. **Target ledger.** Record each applicable country target, voluntary adjustment, baseline, effective month, exemption, and compensation obligation, with document and page identifiers.
2. **Production ledger.** Store both direct-communication and secondary-source estimates where available, with revision vintages and measurement uncertainty.
3. **Capacity and cost intervals.** Use ranges rather than point values, and distinguish sustainable capacity from short-lived surge capacity.
4. **Market controls.** Include global demand indicators, non-OPEC and OPEC+ supply, commercial inventories, freight

and quality differentials, sanctions, outages, and major geopolitical events.

5. **Structural and reduced-form separation.** Estimate descriptive target gaps separately from the structural game. Do not use the same observations both to choose unreported parameters and to validate the model.
6. **Out-of-sample tests.** Fit parameters on an earlier window and evaluate target-gap and price implications on later months. Compare against naive, competitive-fringe, dominant-firm, and no-strategy baselines.
7. **Simulation uncertainty.** For learning agents, report distributions across seeds, convergence diagnostics, policy stability, and sensitivity to reward design. A single trajectory is not evidence.
8. **Falsification.** State what observations would count against the model: persistent output below a profitable modeled best response despite available capacity; target changes unrelated to predicted bargaining pressure; or policy behavior better explained by outages and fiscal constraints than by strategic response.

This roadmap aligns the AI component with scientific use. The simulation becomes a controlled environment for comparing mechanisms and generating testable predictions, not a decorative layer added to a weak dataset.

11 Conclusion

Nash equilibrium explains an important but limited part of OPEC production coordination. It formalizes the temptation facing a producer that can expand output while others remain restrained, and it shows why a target profile must be supported by more than collective interest. Once capacity constraints are imposed and deviations are calculated as conditional best responses, the incentives are both economically intelligible and heterogeneous.

In the illustrative 2023 benchmark, every modeled producer has a positive one-period gain from an upward deviation, so the reconstructed target profile is not a one-shot Nash equilibrium. That result is not the end of the analysis. Repeated interaction can support some countries' targets under plausible patience, but the same punishment is ineffective for others. Monitoring, compensation, tailored continuation payoffs, capacity, and institutional bargaining are therefore central. A statement such as "OPEC is a prisoner's dilemma" is too coarse unless it identifies the players, quantities, payoffs, information, and punishment mechanism.

The multi-agent extension should preserve that discipline. Fixed strategies are useful when labeled as fixed strategies. Reinforcement learning is useful when its reward, state, exploration, non-stationarity, and validation are exposed. The most defensible role for AI here is not to announce a forecast of sovereign behavior. It is to make competing assumptions executable, inspectable, and falsifiable.

The strongest conclusion is consequently methodological: game theory provides a transparent language for OPEC's strategic tensions, while explainable multi-agent computation provides a laboratory for testing how those tensions change under demand shocks, monitoring error, capacity constraints, and enforcement. Neither substitutes for audited data. Used together, they can turn an intuitive story about cooperation and deviation into a cumulative research program.

A Derivations and equilibrium checks

A.1 Interior asymmetric Cournot solution

For an interior equilibrium, the first-order conditions are

$$\alpha - c_i - \beta Q_{-i} - 2\beta q_i = 0. \tag{28}$$

Since $Q_{-i} = Q - q_i$, this becomes

$$\alpha - c_i - \beta Q - \beta q_i = 0, \tag{29}$$

so

$$q_i = \frac{\alpha - c_i - \beta Q}{\beta}. \tag{30}$$

Summing over i gives

$$Q = \frac{n\alpha - \sum_i c_i - n\beta Q}{\beta}, \tag{31}$$

and therefore

$$Q^* = \frac{n\alpha - \sum_i c_i}{(n+1)\beta}. \tag{32}$$

Substitution yields equation (9). Each component must then be checked against $0 < q_i^* < K_i$. If any check fails, the interior formula does not solve the constrained game.

A.2 Complementarity form

Define the marginal-profit expression

$$m_i(q) = \alpha - c_i - \beta \sum_{j \neq i} q_j - 2\beta q_i. \tag{33}$$

The constrained optimum satisfies

$$q_i = 0 \implies m_i(q) \leq 0, \tag{34}$$
$$0 < q_i < K_i \implies m_i(q) = 0, \tag{35}$$
$$q_i = K_i \implies m_i(q) \geq 0. \tag{36}$$

The baseline output from the reproducibility script reports $m_i(K) > 0$ for all nine countries, verifying the all-capacity equilibrium rather than merely assuming it.

A.3 Repeated-game threshold

The no-deviation condition $V_i^C \geq V_i^D$ is

$$\frac{\pi_i^C}{1-\delta_i} \geq \pi_i^D + \frac{\delta_i \pi_i^P}{1-\delta_i}. \tag{37}$$

Multiplying by $1 - \delta_i > 0$ and rearranging gives

$$\delta_i(\pi_i^D - \pi_i^P) \geq \pi_i^D - \pi_i^C. \tag{38}$$

When $\pi_i^D > \pi_i^P$, division yields equation (19). A threshold above one means that even an arbitrarily patient player prefers the deviation path because the designated punishment does not lower its continuation payoff enough.

A.4 Expected marginal enforcement

For $q_i > t_i$, the penalized payoff is

$$\widetilde{\pi}_i(q_i, T_{-i}) = \pi_i(q_i, T_{-i}) - g_i(q_i - t_i). \tag{39}$$

Its right derivative at t_i is

$$\left.\frac{\partial \widetilde{\pi}_i}{\partial q_i}\right|_{t_i^+} = \alpha - c_i - \beta T_{-i} - 2\beta t_i - g_i. \tag{40}$$

Because $\widetilde{\pi}_i$ is concave above the target, the derivative is non-positive for all upward choices exactly when g_i satisfies equation (21).

B Reproducibility package

The accompanying project folder contains:

- `analysis.py`, which transcribes the supplied panel, generates descriptive statistics, calibrates the model, computes equilibrium and repeated-game quantities, runs the sensitivity grid, and creates all figures;
- `tables/panel_2017_2023.csv`, the transcribed source panel, including 2017 for provenance even though the common analysis begins in 2018;
- `tables/descriptive_summary_2018_2023.csv`;
- `tables/model_inputs.csv`;
- `tables/baseline_game_results.csv`;
- `tables/sensitivity_results.csv`; and
- vector PDF and PNG versions of every figure.

The script uses quantities in mb/d, prices and costs in USD per barrel, and profits in million USD per day. No stochastic simulation result is presented in the article because the source model's agents were fixed-rule agents and because a learning model requires additional calibration and validation choices. The architecture in section 7 is therefore a reproducible specification for the next stage rather than a disguised numerical claim.

C Figure descriptions

Figure 1a. A nine-by-six matrix of production-to-target ratios for 2018–2023. Iraq and Gabon are mostly above one. Algeria and Equatorial Guinea are mostly below one. Nigeria and Congo have the largest shifts across years.

Figure 1b. A scatter plot of each country's mean ratio against its standard deviation. The United Arab Emirates lies almost exactly at a mean of one with low variability. Congo and Nigeria have the highest variability.

Figure 2. Grouped horizontal bars compare 2023 target, actual production, and capacity proxy. Saudi Arabia is largest by a wide margin; Iraq is second; small producers' capacities remain below 0.4 mb/d.

Figure 3a. Horizontal bars display one-shot deviation gains. Iraq and Saudi Arabia dominate in absolute dollars; all nine gains are positive.

Figure 3b. Horizontal bars display minimum discount factors. Three countries exceed the feasibility boundary at one, and Congo lies just below it.

Figure 4. Three declining lines show that the constrained-equilibrium price falls as the capacity multiplier rises. The decline is steepest when demand is least elastic.

Contributor, data, and disclosure statements

Contributor roles. Ali Masomzadeh developed the original project, compiled the initial data, and produced the Wolfram Language prototype and first draft. Vandit Rikhi contributed to the original project. Jonathan Kenigson contributed mathematical and methodological review and manuscript revision. All named contributors should review and approve this statement before submission.

Data provenance. The country target, production, and marginal-cost inputs originate in the supplied working draft (Masomzadeh 2026). The reconstruction preserves those values for transparency while labeling target and cost inputs as provisional. The 2023 Brent price and global oil-demand reference values are sourced from EIA and OPEC, respectively (Organization of the Petroleum Exporting Countries 2024b; U.S. Energy Information Administration 2026).

Code and figures. All figures in this manuscript are original computational graphics generated from the accompanying code and tables. They are not generative-AI images.

Generative-AI assistance. Generative AI was used under human direction as an editorial, mathematical-checking, coding, and typesetting aid in reconstructing this manuscript. The named human contributors remain responsible for verifying the data, claims, originality, contributor consent, and final submitted text.

Funding and conflicts of interest. These declarations must be completed and approved by the named contributors before submission. No funding or conflict information was supplied with the source draft.

References

Behar, A. and R. A. Ritz (2017). "OPEC vs US Shale: Analyzing the Shift to a Market-Share Strategy". In: *Energy Economics* 63, pp. 185–198. DOI: `10.1016/j.eneco.2016.12.021`.

Brémond, V., E. Hache, and V. Mignon (2012). "Does OPEC Still Exist as a Cartel? An Empirical Investigation". In: *Energy Economics* 34.1, pp. 125–131. DOI: `10.1016/j.eneco.2011.03.010`.

Caldara, D., M. Cavallo, and M. Iacoviello (2019). "Oil Price Elasticities and Oil Price Fluctuations". In: *Journal of Monetary Economics* 103, pp. 1–20. DOI: `10.1016/j.jmoneco.2018.08.004`.

Cournot, A. A. (1897). *Researches into the Mathematical Principles of the Theory of Wealth*. Trans. by N. T. Bacon. Original French edition published 1838. New York: Macmillan.

Fudenberg, D. and E. Maskin (1986). "The Folk Theorem in Repeated Games with Discounting or with Incomplete Information". In: *Econometrica* 54.3, pp. 533–554. DOI: `10.2307/1911307`.

Golombek, R., A. A. Irarrazabal, and L. Ma (2018). "OPEC's Market Power: An Empirical Dominant Firm Model for the Oil Market". In: *Energy Economics* 70, pp. 98–115. DOI: 10.1016/j.eneco.2017.11.009.

Green, E. J. and R. H. Porter (1984). "Noncooperative Collusion under Imperfect Price Information". In: *Econometrica* 52.1, pp. 87–100. DOI: 10.2307/1911462.

Griffin, J. M. (1985). "OPEC Behavior: A Test of Alternative Hypotheses". In: *American Economic Review* 75.5, pp. 954–963.

Griffin, J. M. and W. Xiong (1997). "The Incentive to Cheat: An Empirical Analysis of OPEC". In: *Journal of Law and Economics* 40.2, pp. 289–316. DOI: 10.1086/467374.

Gülen, S. G. (1996). "Is OPEC a Cartel? Evidence from Cointegration and Causality Tests". In: *The Energy Journal* 17.2, pp. 43–57. DOI: 10.5547/ISSN0195-6574-EJ-Vol17-No2-3.

Kilian, L. and D. P. Murphy (2014). "The Role of Inventories and Speculative Trading in the Global Market for Crude Oil". In: *Journal of Applied Econometrics* 29.3, pp. 454–478. DOI: 10.1002/jae.2322.

Masomzadeh, A. (2026). "Modeling OPEC Quota Compliance Through Game Theory and Nash Equilibrium Analysis". Unpublished Wolfram Language notebook export and working draft supplied by the author.

Nash, J. F. (1950). "Equilibrium Points in n-Person Games". In: *Proceedings of the National Academy of Sciences* 36.1, pp. 48–49. DOI: 10.1073/pnas.36.1.48.

– (1951). "Non-Cooperative Games". In: *Annals of Mathematics* 54.2, pp. 286–295. DOI: 10.2307/1969529.

Okullo, S. J. and F. Reynès (2016). "Imperfect Cartelization in OPEC". In: *Energy Economics* 60, pp. 333–344. DOI: 10.1016/j.eneco.2016.10.010.

Organization of the Petroleum Exporting Countries (Apr. 3, 2024a). *53rd Meeting of the Joint Ministerial Monitoring Committee*. URL: https://www.opec.org/pr-detail/7-03-apr-2024.html (visited on 08/19/2026).

– (2024b). *Annual Statistical Bulletin, 59th Edition*. URL: https://publications.opec.org/asb/archive/123 (visited on 08/19/2026).

– (2026). *Member Countries*. URL: https://www.opec.org/member-countries.html (visited on 08/19/2026).

Smith, J. L. (2005). "Inscrutable OPEC? Behavioral Tests of the Cartel Hypothesis". In: *The Energy Journal* 26.1, pp. 51–82. DOI: 10.5547/ISSN0195-6574-EJ-Vol26-No1-3.

Sutton, R. S. and A. G. Barto (2018). *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge, MA: MIT Press.

Tesfatsion, L. (2003). "Agent-Based Computational Economics: Modeling Economies as Complex Adaptive Systems". In: *Information Sciences* 149.4, pp. 262–268. DOI: 10.1016/S0020-0255(02)00280-3.

U.S. Energy Information Administration (2026). *Spot Prices for Crude Oil and Petroleum Products: Brent—Europe*. Annual averages calculated from daily closing spot prices. URL: https://www.eia.gov/dnav/pet/pet_pri_spt_s1_a.htm (visited on 08/19/2026).

Watkins, C. J. C. H. and P. Dayan (1992). "Q-Learning". In: *Machine Learning* 8, pp. 279–292. DOI: 10.1007/BF00992698.

Wood, A. D., C. F. Mason, and D. Finnoff (2016). "OPEC, the Seven Sisters, and Oil Market Dominance: An Evolutionary Game Theory and Agent-Based Modeling Approach". In: *Journal of Economic Behavior & Organization* 132, pp. 66–78. DOI: 10.1016/j.jebo.2016.06.011.