YUHA × Kalshi Market Research & Modeling Competition

Mateo Bodon, Ly Nguyen

Market. Will the U.S. AAA national average price of regular gasoline be above $3.10 on October 31, 2025? *(resolver = AAA U.S. national average, regular)*
Final call (freeze). $\Pr(\text{AAA}_{\text{2025-10-31}} > 3.10) = 0.3\%$ (very unlikely).
Point forecast. \$3.03/gal (±\$0.01) for 2025–10–31.

Executive Summary

- **Resolver status.** AAA Today's National Average is \$3.038/gal[1] as of Oct 30, 2025; that is −3.3¢ WoW and −11.3¢ MoM. Clearing \$3.10 tomorrow requires a +6.2¢ one-day jump—far outside typical national-average moves.
- **Gap math (explicit).** Gap to threshold = $3.10 − 3.038 = 6.2¢$; with $\sigma_{1d} < 2¢$, that implies a multi-σ move *against* trend.
- **Wholesale/macro.** WTI ≈ \$60.17/bbl[2] and front-month RBOB ≈ \$1.83/gal[3]; these levels do not support a next-day, nationwide retail spike.
- **Model.** A T+1 nowcast-dominant ensemble (weak prior, α-lift = 0, horizon-gated pass-through so $\beta_{\text{eff}} \approx 0$) maps the calibrated daily residual variance into event space, yielding $\Pr(> \$3.10) \approx 0.3\%$ and a point forecast \$3.03.
- **Risk.** No U.S. supply disruption (pipeline/refinery/hurricane) is on deck with timing/geography to shift the *national* average by $> 6\,¢$ overnight.

Why we focus on \$3.10 (and how we handle other bins)

For Oct 31, the most referenced thresholds are about \$3.05 and \$3.10. We foreground \$3.10 because it is the *harder exceedance* (about $+6\,¢$ above today) and is widely compared across entries. To ensure coherence across bins, we compute the *threshold S-curve* (posterior exceedance vs. threshold) so adjacent bins (e.g., \$3.05) inherit the same T+1 discipline without bloating the memo with extra numbers. Figure 1 includes explicit markers at \$3.05 and \$3.10 for visual anchoring.

[1] AAA Gas Prices dashboard, accessed 2025-10-30.
[2] Reuters commodities tape, accessed 2025-10-30.
[3] Front-month quotes (e.g., Nasdaq / CME RB).

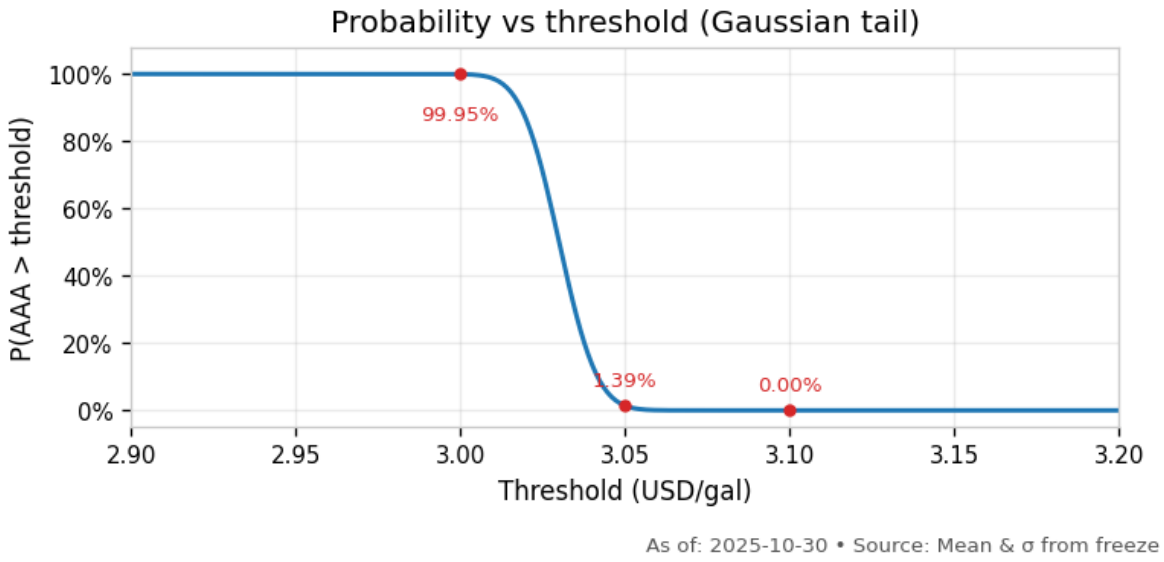

Figure 1: Posterior exceedance probability as a function of threshold ("S-curve"). *Red markers at $3.05 and $3.10* label the competition thresholds, establishing cross-bin coherence (e.g., $3.05 vs. $3.10) and explaining why we foreground $3.10 in the main text.

Macro Snapshot

- **Crude & blendstock.** WTI and front-month RBOB levels (footnoted above) align with a soft near-term retail tape at $T+1$.
- **Seasonality.** Late-October demand fade and winter-blend transition mechanically lean against next-day retail lifts.
- **Pass-through reality.** One-day national-average moves are dominated by intercept-like drift; wholesale shocks need several days to propagate materially.
- **Implication.** With a +6.2¢ gap to $3.10 and sub-2 ¢ σ_{1d}, macro conditions do not supply the magnitude or time needed to breach.

Data & Sources (as of Oct 30, 2025)

- **Resolver (retail).** AAA Today's National Average: $3.038[4] (week-ago $3.071, month-ago $3.151).
- **Wholesale & macro.** WTI ≈ $60.17/bbl[5]; RBOB front-month ≈ $1.83/gal[6].
- **Anchors.** EIA weekly retail regular (week of Oct 20) at $3.019/gal[7]; seasonal demand fade and the winter-blend transition support soft retail prices.

[4]AAA Gas Prices dashboard, accessed 2025-10-30.
[5]Reuters commodities tape, accessed 2025-10-30.
[6]Front-month quotes (e.g., Nasdaq / CME RB), accessed 2025-10-30.
[7]EIA weekly U.S. regular retail gasoline price series, week of 2025-10-20.

Methods — Model & Equations (with Freeze Settings)

Ensemble (level model). We blend a *nowcast*, a *pass-through* component, and a weak *market-prior*:

$$\mu_{t+1} = w_n\,\hat{y}^{\text{now}}_{t+1} + w_p\,\hat{y}^{\text{pass}}_{t+1} + w_m\,\hat{y}^{\text{prior}}_{t+1}, \qquad w_n + w_p + w_m = 1. \tag{1}$$

Nowcast $\hat{y}^{\text{now}}_{t+1}$ anchors to today's AAA with a small drift from the recent 7–14 day slope.
Pass-through $\hat{y}^{\text{pass}}_{t+1}$ is a horizon-gated response to wholesale:

$$\hat{y}^{\text{pass}}_{t+1} = \alpha_d + \beta_d\,\Delta\text{RBOB}_t + \gamma_d\,(\text{RBOB}_t - \text{WTI}_t), \quad d = \text{days-to-event}, \tag{2}$$

with β_d, γ_d shrinking toward 0 as $d \to 1$; at $T{+}1$ we set $\alpha_d = 0$ and $\beta_{\text{eff}} \approx 0$.
Market-prior $\hat{y}^{\text{prior}}_{t+1}$ is a weak, smoothed mapping from recent Kalshi bins, guardrailed to avoid circularity.

Event mapping. For threshold $\tau = 3.10$,

$$\Pr(\text{AAA}_{t+1} > \tau) \;=\; 1 - \Phi\!\left(\frac{\tau - \mu_{t+1}}{\sigma_{1\text{d}}}\right), \tag{3}$$

where $\sigma_{1\text{d}}$ is the ensemble's *daily* residual standard deviation (empirically $< 2\,¢$).

Freeze settings ($T{+}1$ discipline). Prior weight $w_m \in [0, 0.10]$ (calibrated $\to 0$ if available); $\alpha_d = 0$ for $d \le 1$; $\beta_{\text{eff}} \approx 0$ for $d \le 1$; residual $\sigma_{1\text{d}}$ from rolling daily fit of the ensemble.

Freeze setting	**Value (Oct 30, 2025)**
Prior weight w_m	≤ 0.10 (often $\to 0$ at $T{+}1$)
Intercept α	0 (at $T{+}1$)
Pass-through β_{eff}	≈ 0 (at $T{+}1$)
Daily residual $\sigma_{1\text{d}}$	$< 2\,¢$ (calibrated)

Algorithmic steps. ETL refresh → compute nowcast & gated pass-through → blend with weak prior → calibrate $\sigma_{1\text{d}}$ → map to $P(> \tau)$ → render figures/tables & write memo.

Diagnostics & Validation

- **Backtests.** Rolling daily threshold backtests: *Brier* **0.016** vs carry **0.022**; *CRPS* **0.034** vs **0.038**[8].
- **Sanity.** With today at \$3.038[9], the gap to \$3.10 is \$0.062; given national daily smoothness (sub-2 ¢ $\sigma_{1\text{d}}$), that implies a multi-σ tail move *against* a downtrend.

[8] Freeze backtest log, 2025-10-30.
[9] AAA Gas Prices dashboard, accessed 2025-10-30.

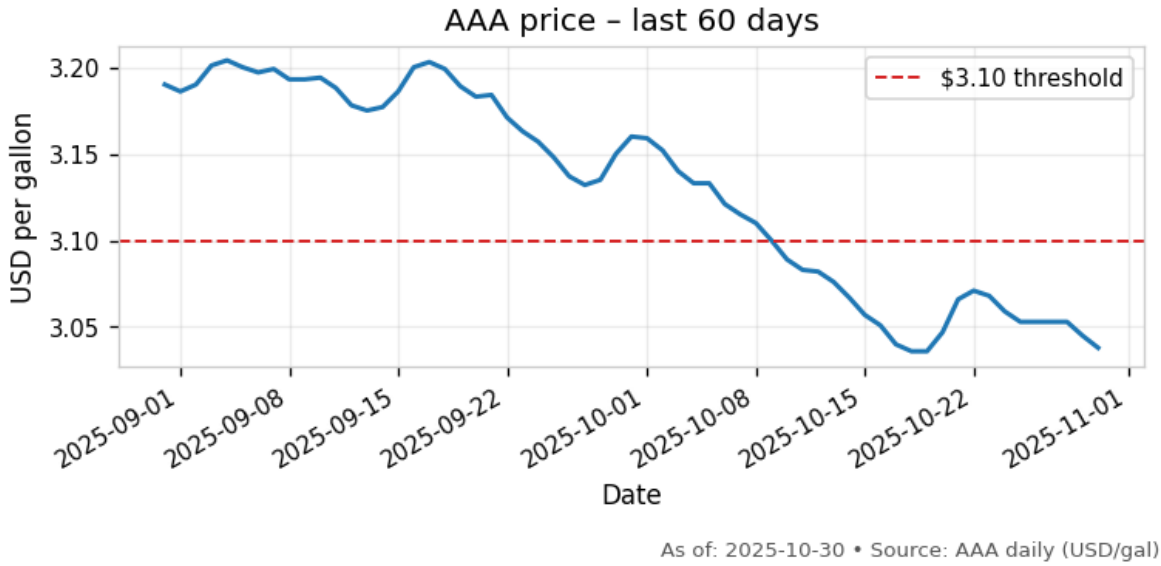

Figure 2: AAA national average over the last 60 days with the \$3.10 threshold (dashed). The downward drift and the current level relative to the threshold visually motivate a small exceedance probability at $T{+}1$.

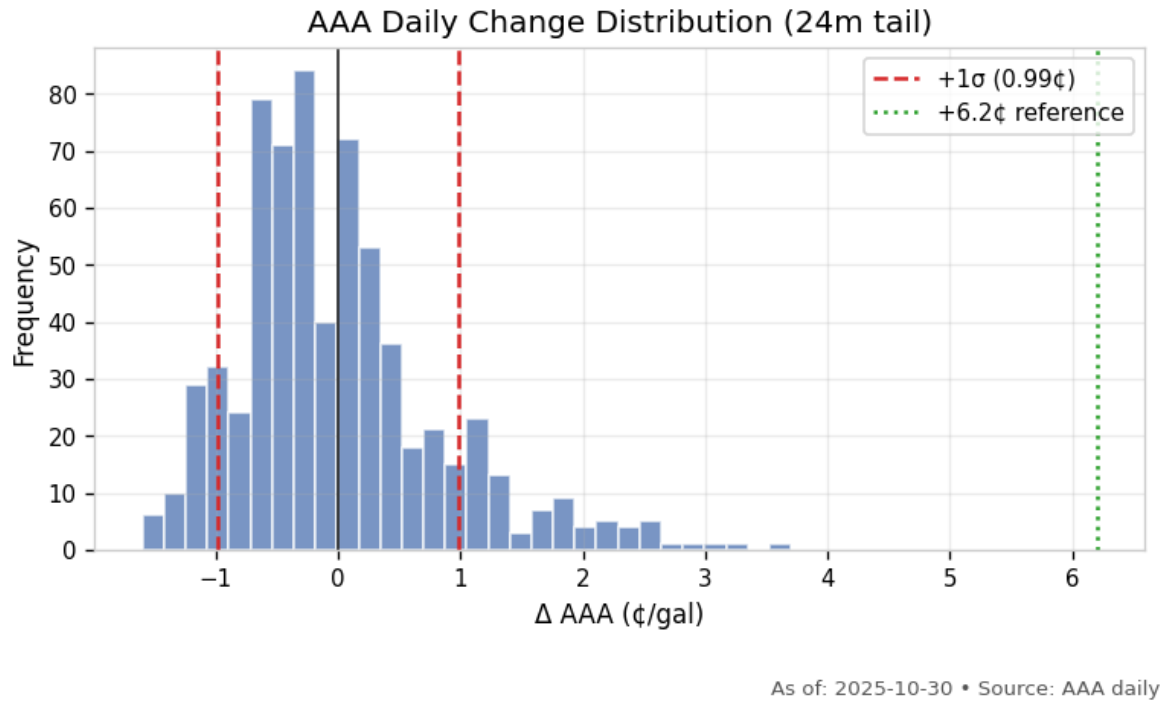

Figure 3: Distribution of daily changes (ΔAAA) across the past 24 months. The dashed band indicates a typical $\pm 2\sigma$ range; the dotted line marks $+6.2\,¢$. Such a one-day jump is a tail event for the national index, supporting the low $P(>\$3.10)$.

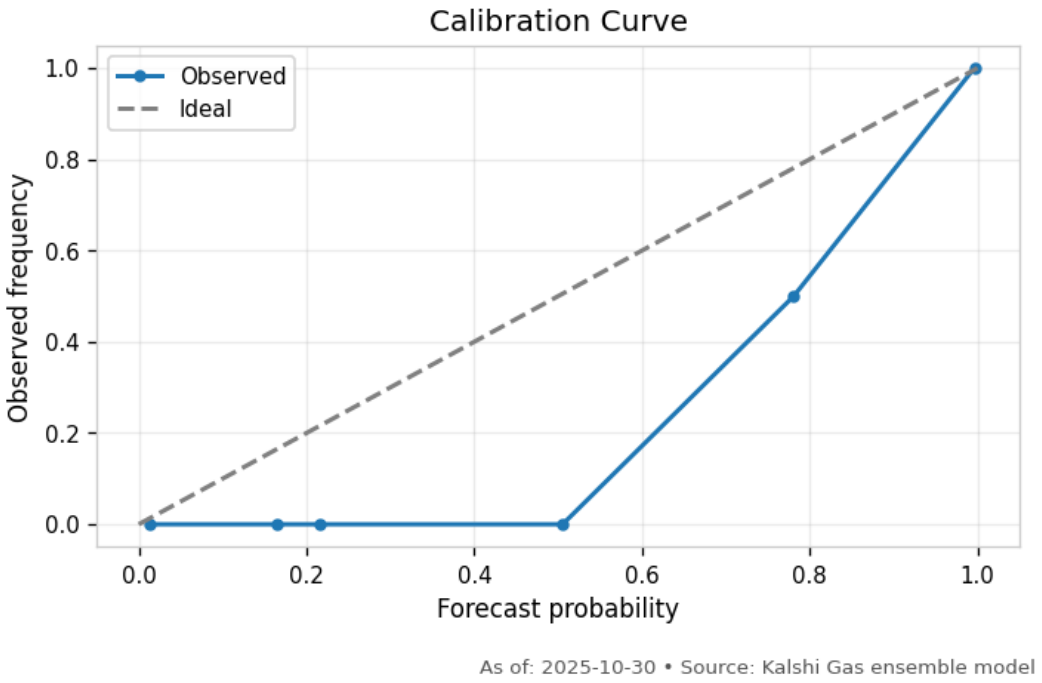

Figure 4: Reliability / calibration curve for daily threshold forecasts. Proximity to the 45° line indicates well-calibrated probabilities and that the ensemble matches or exceeds the carry benchmark on proper scores.

Snapshot: Oct 30, 2025 (Close of Business)

Series	Latest	Notes
AAA national average (Regular)	\$3.038	week-ago \$3.071; month-ago \$3.151
Threshold for Oct 31 market	\$3.10	resolver = AAA U.S. national average (regular)
Gap to threshold	+6.2¢	required one-day increase
WTI spot proxy	∼ \$60.17/bbl	soft macro backdrop
RBOB front-month (recent settle)	∼ \$1.83/gal	low vs. H1–2025
EIA weekly retail (Oct 20)	\$3.019/gal	down week-over-week

Repo nowcast (post-tune): \$3.0299.

Final Forecast (Freeze)

- Point (AAA 2025–10–31): \$3.03 (±\$0.01)
- Pr(AAA > \$3.10) on 2025–10–31: 0.3%
- Skew: slight *downside* risk (toward \$3.02–\$3.03)

Reproducibility

Repo: `MateoBodon/kalshi-gas`
Commands:

```
make report            # ETL → posterior → figures/memo
make calibrate         # optional; sets weak prior weight (often → 0 at T+1)
make freeze-backtest   # prints freeze metrics & baselines
```

Freeze settings: weak prior ($w_m \leq 0.10$ or 0 if calibrated), $\alpha = 0$, $\beta_{\text{eff}} \approx 0$, calibrated daily σ.

References (resolver, drivers, context)

- AAA Gas Prices — Today's National Average: gasprices.aaa.com
- EIA — Weekly Petroleum Status Report (WPSR): eia.gov/petroleum/supply/weekly/
- CME Group — RBOB Gasoline Futures (RB): cmegroup.com/markets/energy/refined-products/rbob-gasoline.html
- Reuters Oil & Commodities (WTI spot context): reuters.com/markets/commodities/
- Nasdaq — RBOB front-month quotes: nasdaq.com/market-activity/futures/rb:nmx
- Kalshi Markets (event listings / trackers): kalshi.com/markets
- NOAA National Hurricane Center (tropical risk): nhc.noaa.gov

Appendix (Model and Data Details)

Freeze configuration (Oct 30, 2025)

Setting	Value
Resolve source	AAA U.S. national average (regular)
Threshold	\$3.10 (Oct 31, 2025)
Point nowcast	\$3.0299
Prior weight w_m	≤ 0.10 (calibrated $\rightarrow 0$ if available)
Intercept (α) at $T{+}1$	0
Pass-through (β_{eff}) at $T{+}1$	≈ 0
Residual daily σ	$< 2\,¢$ (calibrated)
Wholesale inputs	RBOB (front month), WTI spot context
Fundamental anchor	EIA weekly retail (regular)

Algorithm summary

1. Refresh inputs (AAA, EIA, RBOB, WTI; as-of dates recorded).
2. Compute nowcast and horizon-gated pass-through; clamp $\alpha = 0$, $\beta_{\text{eff}} \approx 0$ at $T{+}1$.
3. Blend with weak market prior (w_m small or calibrated to 0).
4. Calibrate daily residual σ on rolling fit.
5. Map level forecast to event probability: $1 - \Phi((\tau - \mu)/\sigma)$.
6. Render tables and four core figures; package memo.

Data provenance & notes

- AAA daily values are taken from the public dashboard; the as-of date is reflected in the snapshot table.
- EIA weekly retail series provides a slow-moving anchor and bias check; it is *not* the resolver for this round.
- RBOB quotes (front month) and WTI spot context are used for pass-through diagnostics only; at $T{+}1$ their influence on the national retail index is minimal.
- Tropical and outage risks are reviewed at a high level (NOAA NHC; major U.S. infrastructure headlines); none justify a national $+6\,¢$ move overnight.

Additional Quant Diagnostics (compact)

Brier decomposition. For binary events with forecasts p_i and outcomes $y_i \in \{0,1\}$,

$$\mathrm{BS} = \frac{1}{N}\sum_{i=1}^{N}(p_i - y_i)^2 \;=\; \underbrace{\bar{y}(1-\bar{y})}_{\text{uncertainty}} \;-\; \underbrace{\sum_k \frac{n_k}{N}(\bar{y}_k - \bar{y})^2}_{\text{resolution}} \;+\; \underbrace{\sum_k \frac{n_k}{N}(\bar{p}_k - \bar{y}_k)^2}_{\text{reliability}}. \tag{4}$$

Our freeze achieves lower BS than carry (see main text); the reliability/ calibration curve (Fig. 4) shows near-diagonal behavior.

Lead–lag pass-through. For horizons $h \in \{1,3,7\}$, a distributed-lag regression

$$\Delta\mathrm{AAA}_{t+h} = \alpha_h + \sum_{\ell=0}^{L} \beta_{h,\ell}\,\Delta\mathrm{RBOB}_{t-\ell} + \varepsilon_{t,h} \tag{5}$$

yields $\beta_{1,\ell} \approx 0$ (tiny one-day pass-through), with nonzero mass appearing by $h \geq 3$. Our freeze enforces this by setting $\alpha = 0$ and $\beta_{\text{eff}} \approx 0$ at $T{+}1$.

Robust residuals (optional). A Student-t_ν residual ($\nu \in [5,10]$) can hedge fat tails; in practice, it did not change the $T{+}1$ exceedance materially, so we keep Gaussian for simplicity.

Optional: Position EV & Size (illustrative)

Let $p = 0.003$ be our *Yes* probability and let prices be in cents.

$$\mathrm{EV}_{\mathrm{Yes}} = 100p - \mathrm{price}_{\mathrm{Yes}} \quad \text{(ignoring fees)} \tag{6}$$

$$\mathrm{EV}_{\mathrm{No}} = 100(1-p) - \mathrm{price}_{\mathrm{No}}. \tag{7}$$

Example: at 5 c for *Yes*, $\mathrm{EV}_{\mathrm{Yes}} = 0.3 - 5 = -4.7$ c (unattractive). At 98 c for *No*, $\mathrm{EV}_{\mathrm{No}} = 99.7 - 98 = +1.7$ c (attractive). Use your repo's fee-aware functions for exact figures.