Forecast Tool Framework

Saadinur

October 24, 2025

1 Mathematical Framework for Forecasting Participation–Issuance Dynamics

1.1 Definitions

We denote discrete protocol rounds by $t = 0, 1, 2, \ldots$, approximately one per day. The key endogenous and exogenous variables are:

$$
\begin{aligned}
P_t \in [0,1] \qquad & \text{Participation rate (fraction of LPT bonded)} \\
P^* \in (0,1) \qquad & \text{Target participation rate (e.g. 50\%)} \\
\gamma_t \geq 0 \qquad & \text{Issuance or inflation rate per round} \\
\mathbf{x}_t \in \mathbb{R}^k \qquad & \text{Market drivers (BTC/ETH/LPT returns, F\&G index etc.)}
\end{aligned}
$$

Policy parameters:

- Inflation change (learning rate): σ

- Issuance bounds: γ_-, γ_+

- Target: P^*

1.2 Dynamics

Protocol policy is:

$$
\gamma_{t+1} = \text{clip}(\gamma_t + \sigma \cdot \text{sgn}(P^* - P_t), \ [\gamma_-, \gamma_+]), \tag{1}
$$

where $\sigma > 0$ controls the per-round step size and $[\gamma_-, \gamma_+]$ are bounds.

To model the evolution of P_t, we work in the log-odds (logit) domain:

$$
Y_t = \text{logit}(P_t) = \ln \frac{P_t}{1 - P_t}.
$$

The dynamics follow a linear autoregressive model with exogenous regressors (Logit–ARX):

$$
Y_{t+1} = \alpha + \phi Y_t + \beta_\gamma \gamma_t + \mathbf{x}_t^\top \beta + \eta_t + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim \mathcal{N}(0, \sigma_\varepsilon^2), \tag{2}
$$

where η_t optionally represents a latent "risk-on / risk-off" sentiment process (AR(1) or Markov switching) (Probably skip at the moment). Given P_t, update Y_{t+1} via protocol policy, test $(\sigma, P^*, \gamma_-, \gamma_+)$. The participation rate is then recovered as

$$
\widehat{P}_{t+1} = \text{logit}^{-1}(Y_{t+1}) = \frac{1}{1 + e^{-Y_{t+1}}}
$$

1.3 Estimation

1. Collect per-round data $(P_t, \gamma_t, \mathbf{x}_t)$.

2. Constrain $\phi \in (0, 1)$ for stationarity (?).

3. Fit:

   - Option A (do this): Estimate parameters $\alpha, \phi, \beta_\gamma, \boldsymbol{\beta}$ using penalized maximum likelihood (ridge regression):

   $$\hat{\boldsymbol{\theta}} = \arg\min_\theta \|Y_{t+1} - (\alpha + \phi Y_t + \beta_\gamma \gamma_t + \mathbf{x}_t^\top \beta)\|^2 + \lambda \|\beta\|^2.$$

   - Option B: Markov Chain Monte Carlo or Variational Inference? to get posterior over $(\alpha, \phi, \beta_\gamma, \beta)$

1.4 Simulation

We simulate paths $\{P_{t+h}, \gamma_{t+h}\}_{h=1}^H$ under chosen policy:

1. Initialize with observed (P_t, γ_t).

2. For each step $h = 1, \ldots, H$ ($H = 180$, 6 months):

$$\gamma_{t+h} = f_{\text{policy}}(P_{t+h-1}; \sigma, P^*, \gamma_-, \gamma_+),$$
$$\epsilon_{t+h} \sim \mathcal{N}(0, \sigma_\epsilon^2)$$
$$\mathbf{x}_{t+h} \sim \text{historical bootstrap}$$
$$Y_{t+h} = \alpha + \phi Y_{t+h-1} + \beta_\gamma \gamma_{t+h-1} + \mathbf{x}_{t+h-1}^\top \beta + \varepsilon_{t+h},$$
$$P_{t+h} = \text{logit}^{-1}(Y_{t+h}).$$

3. Repeat for N paths.

1.5 Forecast Metrics

From simulated paths we compute:

- **Expected time outside bands:**

$$\mathbb{E}_\theta[\#\{h \leq H : P_{t+h} \notin \mathcal{D}_0 = [P_{\text{low}}, P_{\text{high}}]\}]$$

  and analogous metric for issuance γ_t.

- **Quantiles** of time outside bands.

- **Tightest confidence bands** $[L, U]$ at q such that

$$P(L \leq P_{t+h} \leq U, \ \forall h \leq H) \geq q.$$

1.6 Objective and Optimization

Using stakeholder-elicited acceptable bands $\mathcal{D}_0$ for P and γ, define risk-admissibility:

Define a policy parameter vector $\theta = (\sigma, P_\tau, \gamma_-, \gamma_+)$. A configuration is *risk-admissible* if

$$\mathbb{E}_\theta[\text{time outside } \mathcal{D}_0] \leq T^* \quad \text{and} \quad P_\theta(\text{time outside } \mathcal{D}_0 > T_{\text{tail}}) \leq \epsilon,$$

for stakeholder chosen thresholds $T^*, T_{\text{tail}}, \epsilon$. Among admissible sets, choose those minimizing expected dilution and tail band risk.

1.7 Validation and Deliverables

- Backtesting over historical windows with rolling re-estimation.

- Fan charts of P_t and γ_t.

- Tables of risk metrics per candidate policy.

- Identification of admissible policy sets.

- "Admissible set" finder: grid search (or Bayesian optimization) over θ minimizing expected dilution subject to band constraints. (?)

2 More on Estimation: Step 3 above.

The estimation step quantifies the relationship between current variables and the next-round participation rate. It provides the fitted parameters $\theta = (\alpha, \phi, \beta_\gamma, \beta, \sigma_\varepsilon)$ that govern the transition dynamics of the model.

2.1 Step 1. Data Prep

Let the data be observed for rounds $t = 1, \ldots, T$, including participation rate P_t, issuance rate γ_t, and external variables $\mathbf{x}_t \in \mathbb{R}^k$. First, transform the participation rate into the logit domain:

$$Y_t = \ln \frac{P_t}{1 - P_t}.$$

This ensures $P_t \in [0, 1]$ and linearizes its evolution in the state equation.

2.2 Step 2. Model Setup

The predictive design matrix is constructed such that the current state explains the next step:

$$Y_{t+1} = \alpha + \phi Y_t + \beta_\gamma \gamma_t + \mathbf{x}_t^\top \beta + \varepsilon_{t+1}.$$

Hence, the regressors are $X_t = [1, Y_t, \gamma_t, \mathbf{x}_t]$, and the target variable is Y_{t+1}. Optionally, additional lags $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, \ldots$ can be included to capture persistence in external drivers.

2.3 Step 3. Model Validation

To avoid look-ahead bias, the dataset is partitioned sequentially:

- A *training window* of length T_{train} (e.g., 365 rounds) is used for estimation,

- A *validation window* of T_{val} rounds is used for out-of-sample testing,

- The origin is rolled forward (expanding or sliding window) to produce a series of re-estimated models.

This mirrors real-time use of the model and enables evaluation of forecasting accuracy.

2.4 Step 4. Parameter Estimation

The base estimator is a penalized least squares (ridge regression):

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \left\| Y_{t+1} - (\alpha + \phi Y_t + \beta_\gamma \gamma_t + \mathbf{x}_t^\top \beta) \right\|^2 + \lambda \|\boldsymbol{\beta}\|^2,$$

where $\lambda > 0$ stabilizes coefficients in the presence of multicollinearity. The residual variance σ_ε^2 is estimated from the regression errors.

The coefficients $(\alpha, \phi, \beta_\gamma, \beta)$ describe the conditional mean process for Y_{t+1}, and the estimated variance σ_ε^2 characterizes the innovation noise used in simulations.

2.5 Step 5. Integration with Prior OLS Results

Previous OLS or log–log regressions estimating participation drivers can be reused here:

- Coefficients from the prior OLS provide informed initial values for β_γ and β,

- The new model extends the prior regression by adding:

    1. An autoregressive term Y_t (persistence of participation),
    2. The logit transformation of P_t,
    3. Sequential re-estimation over time to capture evolving dependencies.

- Ridge or Bayesian regularization replaces simple OLS to improve robustness.

Thus, the previously discovered relationships remain valid inputs but are embedded in a dynamic and forecast-oriented framework.

2.6 Step 6. Transition to Simulation Phase

Once parameters are estimated, the fitted model yields a stochastic transition rule:

$$\widehat{Y}_{t+1} = \hat{\alpha} + \hat{\phi} Y_t + \hat{\beta}_\gamma \gamma_t + \mathbf{x}_t^\top \hat{\boldsymbol{\beta}} + \varepsilon_{t+1}, \qquad \varepsilon_{t+1} \sim \mathcal{N}(0, \hat{\sigma}_\varepsilon^2).$$

These quantities serve as inputs for the Monte Carlo simulation of (P_t, γ_t) trajectories under alternative policy parameters.

References

[1] Andrew Macpherson. *Livepeer issuance parameter setting and risk analysis.* 2025. URL: https://hackmd.io/@i79XZRmjR86P6AbhL0jwVQ/Sk4jTF7pxg.