# Dynamic Generator Inversion for Observable Conditional Distributions

Johannes Bleher

June 5, 2026

**Abstract**
Economic time series often record discrete adjustments, while the empirical question concerns the conditional distribution of observables such as prices, quantities, inventories, choices, durations, or stress states. This paper introduces generator inversion as a distributional econometric framework. The estimand is the conditional adjustment generator: the law that maps current states into future events or transitions. The estimator is a validation-selected regularized conditional empirical-characteristic-function (ECF) minimum-distance procedure, optionally anchored by likelihood or transition fitting. The key restriction is that the generator must earn its value through observables. Together with a measurement law, the selected generator induces one predictive distribution from which means, variances, quantiles, intervals, tail probabilities, and rare-event scores are all derived. Vector autoregressions, Gaussian transition models, Hawkes specifications, and quantile regressions then become moment projections, likelihood anchors, score projections, or direct benchmarks for the same forecast problem. Held-out characteristic losses test the adjustment law and the observable forecast transform; proper scoring rules test the resulting observable distributions. The executable simulation is a diagnostic laboratory rather than an empirical market claim. It shows how generator estimates, ECF regularization, observable distribution forecasts, and reduced-form benchmarks can be compared under one validation design. The contribution is to turn event-generator estimation from a formal representation into a practical method for forecasting and testing observable conditional distributions.

**Keywords:** dynamic generators; empirical characteristic functions; conditional distributions; distributional forecasting; quantile regression
**JEL codes:** C14; C32; C51; C53; C58

# 1 Introduction

Economic time series often summarize systems that move through discrete adjustments. Firms revise prices, processors reallocate production, inventories change, limit orders arrive and are cancelled, borrowers enter default, and workers switch jobs. The empirical target is usually not the event history itself. Economists want conditional distributions of observables: prices, quantities, inventories, market shares, durations, or stress states. This paper asks whether those distributions can be modeled better by first estimating the dynamic law of adjustment that generates them.

Existing reduced-form approaches typically model observables directly. A vector autoregression projects conditional means, a volatility model projects second moments, a Hawkes process captures event clustering, a Gaussian transition density imposes a low-order approximation, and quantile regression estimates selected conditional quantiles. These methods are useful because they target observable objects. They also leave open a common problem: different adjustment laws can agree on a mean, variance, or a few quantiles while implying different tail probabilities, event composition, counterfactual paths, and conditional characteristic functions. A method that estimates the adjustment law can therefore be useful only if it improves or disciplines the conditional distributions of observables.

The paper develops generator inversion as a distributional econometric framework. Its goal is to turn generator estimation from a formal representation into a practical forecasting and testing method for observable outcomes. The estimand is the conditional adjustment generator. It maps the current state into the law of future events or transitions. The estimator is an adaptive regularized conditional empirical-characteristic-function (ECF) estimator. It selects, on validation data, a finite characteristic design consisting of frequencies, state instruments, block length, weighting rule, and regularization around a likelihood or transition anchor. The output is not merely an estimated intensity surface. Once the generator and the measurement law are fixed, they induce a full conditional distribution for observable outcomes. That distribution is the object evaluated by adjustment and observable characteristic-function loss, quantile loss, interval scores, tail-probability scores, and rare-event diagnostics.

The contribution is an estimand-to-score chain. First, the paper defines the conditional adjustment generator as an estimand for event-history and transition-panel data. Second, it proposes an adaptive regularized conditional ECF estimator that chooses the characteristic design and regularization

on validation data, including the possibility of staying at a likelihood or transition anchor when characteristic moments do not improve out of sample. Third, it maps the estimated generator into observable outcome distributions through a transition kernel and a measurement or projection law. Fourth, it evaluates those distributions with held-out characteristic-function, quantile, interval, tail, and rare-event scores. Fifth, it interprets standard econometric models as moment projections, likelihood anchors, score projections, or final-test benchmarks of the same observable forecast problem rather than as unrelated alternatives.

The phrase full conditional distribution has a precise scope in this paper. It does not mean that a finite sample and a finite grid of characteristic moments recover an unrestricted nonparametric conditional law without structure. It means that, after the adjustment generator, characteristic design, regularization rule, and measurement law have been selected before final testing, all reported forecast objects are derived from one conditional predictive distribution. Under correct specification and separating characteristic restrictions, this distribution is the target law on the visited support. Under sparse designs, regularization, or misspecification, it is the validation-selected generator projection whose observable consequences can be tested against reduced-form benchmarks.

The main formal claim is the generator-to-observable link. Under the stated fixed-design ECF, validation-selection, and observable-mapping conditions, Proposition 22 shows that the validation-selected generator induces a consistent observable forecast law. Under misspecification or sparse characteristic designs, the same result defines the target as a validation-selected generator projection. The final score gap is therefore a held-out comparison of observable forecast distributions, not a goodness-of-fit statistic for a latent object.

The paper connects four econometric literatures that are usually kept separate. ECF methods provide distributional moment restrictions (Feuerverger and Mureika, 1977; Carrasco and Florens, 2000; Singleton, 2001). Event-history, point-process, and Hawkes models provide likelihood anchors for adjustment data (Andersen and Gill, 1982; Hawkes, 1971). Quantile and forecast-evaluation methods provide observable distributional benchmarks (Koenker and Bassett, 1978; Gneiting and Raftery, 2007). Inversion arguments in demand estimation clarify why recovering a latent economic object can be useful for counterfactuals and market-specific distributions (Berry, 1994; Berry and Haile, 2014; Albrecht and Traina, 2026). The connection is an estimand analogy, not a claim that this paper estimates a demand curve. Static demand inversion recovers a market-specific object before counterfactual quantities are computed. Generator inversion recovers a state-specific law

of adjustment before observable distributions are forecast and scored. The paper's contribution is to combine these elements in one estimand-to-score chain: estimate an adjustment generator, regularize its characteristic restrictions, map it into observable distributions, and compare those distributions with direct reduced-form forecasts.

Table 1: Core objects in generator inversion.

The table summarizes the paper's five-part framework: the estimand, estimator, bridge to observables, evaluation criteria, and interpretation of standard reduced-form models.

| Component | Object | Framework role |
|---|---|---|
| Estimand | Conditional adjustment generator | Law mapping current states into future events or transitions |
| Estimator | Adaptive regularized empirical-characteristic-function minimum distance | Validation-selected characteristic design and regularization against a likelihood or transition anchor |
| Bridge | Observable forecast law | Transition kernel plus measurement or projection law mapping adjustment paths into prices, quantities, states, or other outcomes |
| Evaluation | Distributional scoring rules | Characteristic-function loss for the adjustment law and observable distribution; quantile loss, interval scores, tail scores, and score gaps for observables |
| Interpretation | Reduced-form score projections and benchmarks | Vector autoregressions, Gaussian transition models, Hawkes processes, and quantile regressions as moment projections, score projections, likelihood anchors, or benchmarks |

Table 2 states the claim boundary. The framework is distributional because one fitted adjustment law and measurement law generate all reported forecast summaries. It is not unrestricted: finite characteristic designs, validation regularization, transition-only observation, and misspecification all change the target from a structural generator to a selected forecast projection.

Table 3 introduces the core adjustment-law notation used throughout the paper. The ECF and forecast-evaluation symbols are introduced later, next to the calculations in which they are used, so the notation is tied to the corresponding empirical object rather than collected in one overloaded dictionary. Appendix C then collects the calculation trail in one place for reference.

To keep later calculations readable, the paper uses distinct symbols for objects that are easy to confuse. The letter $K$ denotes the basis dimension in generator estimation, while $K_q$ denotes the number of reported quantile levels. The ECF design complexity price is $\xi$, not a quantile level. The fixed rare-action horizon is $h_{\text{act}}$. The conjugate response field in Hamiltonians and actions is $\varpi$, reserving $p_t$ and $p_j$ for economic prices.

Table 2: Scope of the full-distribution claim.

The table separates what the paper claims from what it deliberately does not claim. It clarifies when the estimated object is a local generator, a transition law, or a validation-selected forecast projection.

| Question | Claim made in the paper | Boundary of the claim |
|---|---|---|
| *Object recovered* | | |
| What is estimated? | A conditional adjustment law on the state support reached by the data. | Transition panels identify an observed-horizon transition law unless additional semigroup, multi-horizon, or injectivity restrictions justify a generator interpretation. |
| What does ECF add? | Characteristic restrictions regularize and test the distributional implications of the adjustment law across selected frequencies and state instruments. | A finite frequency and instrument design identifies a selected projection, not an unrestricted nonparametric law. |
| *Observable distribution* | | |
| What is the full distribution? | All reported forecast objects are functionals of one generator-implied predictive distribution after the measurement law is fixed. | The distribution is true only under correct specification and separating restrictions; otherwise it is the validation-selected generator projection. |
| What is evaluated? | Held-out characteristic losses test the adjustment law; proper scoring rules test the observable forecast distribution. | Adjustment-law fit and observable forecast value are separate claims and may point in different directions. |
| *Benchmark interpretation* | | |
| How do reduced-form models enter? | Vector autoregressions, Gaussian transition models, Hawkes models, probability models, and quantile regressions are moment projections, score projections, likelihood anchors, or final-test benchmarks. | The generator adds value only when adjustment information improves or disciplines observable distributional scores out of sample. |

The statistical estimand is the conditional adjustment generator. The characteristic representation is the main estimation device because it turns the generator into testable restrictions on conditional distributions. For each adjustment type $r \in \mathcal{R}$, let $\nu_r$ be its effect on the state and let $\lambda_{r,\theta}(x)$ be its intensity at state $x$ under generator parameter $\theta$. The Doi–Peliti Hamiltonian evaluated at imaginary momentum,

$$H_\theta(x, \mathrm{i}u) = \sum_{r \in \mathcal{R}} \lambda_{r,\theta}(x)\{\exp(\mathrm{i}u^\top \nu_r) - 1\},$$

is the local conditional characteristic exponent of the aggregate increment. This object keeps the distributional content of the generator visible. Its derivatives at $u = 0$ are conditional cumulants. Its quadratic truncation is the Gaussian diffusion approximation. Its evaluation away from zero gives characteristic-function restrictions for minimum-distance estimation and held-out testing.

The practical estimator uses a basis-function generator. For each event type,

$$\lambda_r(x) = e_r(x) \exp\left\{\sum_{k=1}^{K} \theta_{rk}\phi_k(x)\right\}, \tag{1}$$

where $e_r(x)$ is an exposure term and $\phi_k$ are state functions. The likelihood version is a standard point-process target when events are observed, and an analogous transition likelihood is available

Table 3: Core adjustment-law notation.

The table introduces the main symbols for the adjustment state, observable outcome, event types, transition effects, intensities, generators, transition kernels, and generator parameters.

| Symbol | Name | Role in the framework |
|---|---|---|
| $X_t$ | Adjustment state | State observed before an event or transition; the conditioning object for the generator. |
| $Y_t$ | Observable outcome | Price, quantity, inventory, duration, state, or other outcome whose conditional distribution is evaluated. |
| $\chi$ | Direct observable map | Function that maps the adjustment state into the observable outcome when the outcome is measured directly from the state. |
| $h$ | Forecast horizon | Horizon over which the future state, observable increment, and forecast distribution are evaluated. |
| $M_h(\mathrm{d}y \mid x, x')$ | Measurement law | Distribution that maps a starting state and future adjustment state into an observable increment over horizon $h$. |
| $r \in \mathcal{R}$ | Adjustment type | Event category such as an arrival, removal, revision, transition, or reallocation. |
| $\nu_r$ | State change from event $r$ | Local movement of the adjustment state when event $r$ occurs. |
| $\rho_r^Y(x)$ | Observable change from event $r$ | Event-specific movement of the observable outcome induced at state $x$. |
| $\lambda_r(x)$ | Conditional event intensity | Instantaneous adjustment rate at state $x$ in an event-history model. |
| $A_Y(x)$, $B_Y(x)$ | Observable drift and second-moment rates | Generator-implied local projections for the observable outcome, obtained by aggregating event-specific observable changes with event intensities. |
| $\alpha_h$, $\beta_h$, $\gamma_h$, $\delta_h$, $\varepsilon_{t+h}$ | Projection intercept, loadings, and error | Horizon-specific coefficients and residual in the observable projection that maps generator signals and aggregate controls into an outcome forecast. |
| $w_t$ | Aggregate controls | Observable controls, such as lagged outcomes or aggregate state variables, used alongside generator-implied signals in projection benchmarks. |
| $\mathcal{L}_\theta$ | Conditional adjustment generator | Operator mapping the current state into the local law of future adjustment. |
| $K_{\theta,\Delta}(\mathrm{d}x' \mid x)$ | Transition kernel | Distribution of the future state $x'$ after horizon $\Delta$, conditional on current state $x$. |
| $\theta \in \Theta$ | Generator parameter | Coefficients indexing a candidate generator; $\Theta$ is the parameter space searched by likelihood or ECF estimation. |

when only coarser transitions are observed. The paper's ECF estimator minimizes conditional empirical-characteristic-function moments over a prespecified library of characteristic designs. Validation data select the frequency and instrument design, the block length and weighting rule, and the regularization around the likelihood or transition fit. If the ECF moments do not support moving away from the anchor, the estimator keeps the benchmark generator. If richer characteristic moments improve held-out loss enough to justify their complexity, the estimator moves away from the anchor. This makes the estimator adaptive rather than automatically more complex.

Doi–Peliti methods enter as a representation language, not as the contribution by themselves. The construction was developed for classical Markov jump systems (Doi, 1976; Peliti, 1985). Here it supplies the Hamiltonian that links event intensities to characteristic functions, cumulant expansions, diffusion projections, and large-deviation actions. Familiar econometric models then appear as moment projections, likelihood anchors, score projections, or direct benchmarks. Autoregressive and vector-autoregressive models approximate drift. Conditional-volatility equations approximate second conditional cumulants. Hawkes models capture event clustering and can anchor event-history comparisons. Gaussian transition models set higher conditional cumulants to zero. Quantile regression estimates observable quantiles directly under pinball loss rather than through the generator.

The core claim is deliberately empirical and testable. A dynamic generator is useful not because it is a richer latent object, but because it induces conditional distributions for observables that can be scored against reduced-form benchmarks. Proposition 2 formalizes the resulting value condition: if the benchmark information is sufficient for the observable distribution, the generator should not be expected to improve proper forecast scores; if the adjustment state changes distributional features that the benchmark cannot see, the generator has something to contribute. The paper therefore evaluates the generator at several levels: event likelihood, conditional characteristic-function loss, observable mean and volatility loss, full predictive distributions, quantile loss, prediction intervals, tail probabilities, and rare-state ranking. This design allows the generator to fail where aggregation is sufficient and to add value where adjustment information contains predictive information that lower-order projections discard.

The paper uses a demand–supply event simulation as a diagnostic laboratory, not as the central empirical question. No single market is the contribution. The simulation lets every object in the framework be observed: event intensities, the implied characteristic generator, aggregate projections, observable distribution forecasts, and reduced-form benchmarks. This makes it possible to

check whether the ECF step improves held-out characteristic loss, whether generator signals add value for observable forecasts, and whether direct quantile regression already captures the relevant distributional information. The mixed evidence is central to the paper's discipline: generator inversion must earn its place by improving, organizing, or falsifying observable distributional forecasts.

The demand–supply and market-specific examples should be read as applications of the general framework. In a static demand setting, Berry inversion recovers a latent demand index needed for market-specific counterfactuals (Berry, 1994; Berry and Haile, 2014). The dynamic analogue is to recover the conditional law that governs adjustment before computing observable outcome distributions. In an order book, events are high-frequency arrivals and removals. In industrial or agricultural markets, events may be procurement-price revisions, inventory changes, production reallocations, or benchmark-price transmissions. The econometric object is the same: a conditional law of adjustment whose value is assessed through the observable distributions it implies.

The remainder of the paper is structured as follows. Section 2 states the generator-inversion argument. Section 3 introduces the generator notation. Section 4 defines the characteristic generator and the conditional characteristic-function diagnostic. Section 5 links the generator to observable outcome distributions. Section 6 relates the object to standard econometric projections. Section 7 gives one demand–supply interpretation of generated adjustment curves. Section 8 uses a simulation diagnostic to exercise the estimator, ECF regularization, observable-distribution forecasts, and benchmark scores. Section 9 states identification and sieve-inference arguments. Section 10 defines application protocols for event histories and transition panels. Section 11 concludes. The appendix contains the minimal birth–death derivation, reproducible numerical checks, and notation-to-calculation maps.

## 2 Generator Inversion

Generator inversion starts from a simple distinction. A direct forecast model estimates a feature of an observable distribution. A generator model estimates the conditional law that produces the future adjustment path and then derives observable distributions from that law. The second route is useful only when the estimated generator contains distributional information that the direct forecast discards or when it imposes a coherent restriction across several forecast targets.

Let $X_t$ denote the state of an economic system and let $Y_t = \chi(X_t)$ be an observable outcome when

the outcome is directly measured from the state. More generally, a measurement law maps future states into observables. The adjustment generator is the conditional law of the next local movement in $X_t$. In a jump setting it is a collection of state-dependent event intensities $\lambda = \{\lambda_r\}_{r\in\mathcal{R}}$, where event $r$ changes the state by $\nu_r$. In a coarser transition setting the directly identified object is the conditional transition law for $X_{t+h}$ given $X_t$, where $h$ is the forecast horizon. That transition law supports a generator interpretation only when event histories, semigroup-linked horizons, or an injective transition-to-generator parameterization supply the extra structure. Generator inversion therefore means recovering the conditional adjustment law identified by the observation regime and using it to infer the distribution of $Y_{t+h} - Y_t$.

The method has a fixed sequence. Table 4 states the workflow before the paper develops each component formally. The purpose is to keep the estimator tied to the observable evaluation problem: every step either identifies the adjustment law, regularizes its distributional implications, or translates it into a forecast distribution that can be scored.

Table 4: Generator inversion as an econometric workflow.
The table lays out the sequence from event or transition data to an estimated adjustment law, then from that law to observable distribution forecasts and benchmark comparisons.

| Step | Econometric task | Output used in the next step |
|---|---|---|
| 1 | Define the observation regime and state. | Event histories or transition panels, pre-adjustment states, horizons, and observable outcomes. |
| 2 | Specify the adjustment law. | Event intensities, transition kernels, or generator restrictions that map current states into future adjustment laws. |
| 3 | Fit an anchor adjustment model. | Likelihood, transition-density, or prespecified transition fit used for initialization and benchmarking. |
| 4 | Construct conditional characteristic moments. | Frequency-instrument restrictions comparing realized increments with generator-implied characteristic functions. |
| 5 | Select the regularized ECF estimator. | Validation-selected adjustment law, including the possibility of keeping the anchor when ECF moments do not improve held-out loss. |
| 6 | Translate the adjustment law into observable distributions. | Predictive distributions for prices, quantities, states, durations, or other researcher-facing outcomes. |
| 7 | Evaluate against reduced-form benchmarks. | Characteristic-function losses, quantile losses, interval scores, tail scores, rare-event scores, and benchmark gaps. |

For an application to be credible, the researcher must specify the adjustment object, the observable target, and the benchmark design before final testing. Table 5 collects these requirements. This table is deliberately domain-neutral. In high-frequency data the adjustment object may be an event

stream; in lower-frequency panels it may be a transition record. In both cases, the fitted adjustment law is only useful after a measurement or projection law translates the adjustment distribution into the conditional distribution of an observable outcome.

Table 5: Requirements for applying generator inversion.

The table lists the objects that must be fixed before generator inversion can be estimated and evaluated in an event-history or transition-panel application. Panel headers separate the adjustment object, generator estimation, observable distribution, and final evaluation.

| Requirement | Must be specified | Why it matters |
|---|---|---|
| *Adjustment object* | | |
| Observation regime | Event history or transition panel, including the observation horizon. | Determines whether the anchor is an event likelihood, a transition fit, or a simulated transition law. |
| State before adjustment | Information observed before the event or transition. | Defines the conditioning set for the generator and the instruments for conditional ECF moments. |
| Adjustment increment | State movement over the event or transition horizon. | Supplies the realized object compared with the generator-implied characteristic function. |
| *Generator estimation* | | |
| Candidate generator family | Event intensities, transition kernels, or simulation law indexed by parameters. | Defines the estimand and the parameter space over which the ECF criterion is minimized. |
| Anchor fit | Likelihood, transition density, or prespecified reduced transition model. | Provides initialization, a benchmark, and the center of regularization. |
| ECF design | Frequency bands, state instruments, block length, weighting rule, and regularization grid. | Determines which conditional distributional restrictions are tested and how sampling dependence is handled. |
| *Observable distribution* | | |
| Observable target | Price, quantity, inventory, state, duration, market share, or rare-event outcome. | Defines the economic object whose conditional distribution is scored. |
| Measurement or projection law | Mapping from future adjustment states to the observable outcome. | Turns the estimated generator into a predictive distribution for observables. |
| Counterfactual rule, if used | Intervention on event intensities, transition kernels, measurement laws, information sets, or initial-state distributions. | Separates forecast claims from policy or counterfactual claims and states the required support and invariance assumptions. |
| Reduced-form benchmark library | Vector autoregression, Gaussian transition model, Hawkes projection, quantile regression, probability model, and the rule used to select or retain them before final testing. | Sets the standard that the generator must beat or discipline out of sample without using the final test block for benchmark search. |
| *Evaluation* | | |
| Split rule | Training, validation, and final test blocks fixed before final reporting. | Separates estimation, regularization choice, and evidence used for the paper's claims. |
| Scoring rules | Characteristic loss, quantile loss, interval score, tail probability loss, rare-event ranking, and point-forecast loss. | Evaluates the full conditional distribution rather than only a conditional mean. |
| Interpretation rule | How mixed characteristic and observable-score evidence is reported. | Distinguishes adjustment-process fit from observable forecast value. |

The reporting standard is equally important. A paper using generator inversion should not report only a fitted generator or only an improved forecast score. It should show how the adjustment law is identified, how the ECF restrictions move or retain the anchor, how the fitted law is translated into an observable distribution, and which benchmark conclusions survive final testing. It should also separate executable diagnostics from empirical claims: a simulation can verify the estimator and scoring chain, but only final-test evidence on the target data can support a domain conclusion.

Table 6 states this minimum evidence package. It is meant to make the framework auditable across domains: an order-book application, a processor panel, or a credit-transition panel can use different data, but each must report the same chain from adjustment-law evidence to observable distribution evidence.

Table 6: Minimum evidence package for generator inversion.

The table lists the evidence that should accompany an application of generator inversion. Panel headers separate identification and design, adjustment-law evidence, observable-distribution evidence, and interpretation boundaries.

| Evidence block | What is reported | Claim disciplined |
|---|---|---|
| *Identification and design* | | |
| Observation regime | Event history, transition panel, observed horizon, and support of the conditioning states. | States whether the object is an event generator, a transition law, or a generator claim that needs extra semigroup or injectivity structure. |
| Prespecified design | State variables, adjustment increments, ECF frequencies, instruments, block length, regularization grid, and sample split. | Prevents the characteristic design and benchmark set from being chosen on the final test block. |
| *Adjustment-law evidence* | | |
| Anchor and ECF fit | Anchor fit, selected ECF design, selected regularization, held-out characteristic loss, and loss gap relative to the anchor or benchmark. | Shows whether the estimated adjustment law improves conditional characteristic restrictions beyond the first-stage fit. |
| Stability and uncertainty | Block-bootstrap interval, positive bootstrap share, and sensitivity to admissible block lengths or designs. | Distinguishes robust characteristic evidence from finite-sample movement in a noisy moment set. |
| *Observable-distribution evidence* | | |
| Forecast construction | Observable target, measurement or projection law, simulation size if used, and reduced-form benchmark forecasts. | Makes clear how the generator becomes a predictive distribution for an observed outcome. |
| Benchmark construction | Benchmark family, information set, forecast horizon, and selection rule fixed before final testing. | Ensures score gaps compare forecast distributions for the same target rather than advantaged information sets or ex post benchmark choices. |
| Score comparison | Quantile loss, interval score, tail probability loss, full-distribution score, and point-forecast loss when relevant. | Tests whether the generator improves the observable conditional distribution, not only the adjustment history. |
| *Interpretation* | | |
| Claim boundary | Target, horizon, information set, and scores for which the generator wins, ties, or loses. | Restricts the conclusion to the observable objects supported by final-test evidence. |
| Benchmark conclusion | Whether a reduced-form benchmark, the anchor, or the generator-implied distribution is retained for the reported target. | Turns mixed evidence into a transparent conclusion about predictive information in the adjustment layer. |
| Counterfactual conclusion, if reported | Intervention rule, invariant generator or measurement components, support region, and resulting observable distribution under the counterfactual regime. | Prevents a predictive generator fit from being interpreted as a policy or structural counterfactual without the extra assumptions such claims require. |
| *Reproducibility and scope* | | |
| Executable checks | Estimator tests, normalization checks, table-generation checks, and paper-build status for the reported evidence. | Shows that the estimator, observable forecast construction, and reported tables are reproducible rather than only described. |
| Evidence boundary | Which results are simulation diagnostics, which are empirical final-test results, and which claims remain untested for the target domain. | Prevents simulation recovery or adjustment-law fit from being reported as an observable empirical conclusion. |

The central estimand-to-evaluation link can be stated without committing to a specific market. Fix a horizon $h$, a state $x$, a candidate adjustment generator $\theta$, and a measurement or projection law $M_h$ that maps future adjustment states into the observable outcome. The generator implies a transition

kernel $K_{\theta,h}$ and therefore an observable distribution $F_{Y,\theta}(\cdot \mid x)$. This distribution is the object from which the researcher derives conditional means, variances, quantiles, prediction intervals, tail probabilities, and rare-event probabilities. A reduced-form benchmark may estimate any one of these objects directly. Generator inversion instead asks whether one fitted adjustment law produces a coherent observable distribution that scores well across the targets fixed before final testing.

The word full is used in this conditional sense. The paper does not claim that a finite sample and a finite ECF grid recover an unrestricted nonparametric conditional distribution without structure. The claim is that, once the adjustment generator and the observable mapping have been selected, all reported forecast objects are functionals of one forecast distribution. If the characteristic design separates the true conditional law and the model is correctly specified, this distribution is the true observable law on the visited support. If the design is sparse, regularized, or misspecified, it is the validation-selected generator projection. The final-test scores distinguish these cases empirically by asking whether that projected distribution improves observable forecasts relative to prespecified benchmarks.

Forecast and counterfactual claims use the same mathematical bridge but require different assumptions. A forecast holds the adjustment law and measurement law fixed at the law learned from the observed data. A counterfactual changes some part of that system: an event intensity, a transition kernel, a measurement law, an initial-state distribution, or the information available to agents. Let $a$ denote a prespecified counterfactual regime, with transition kernel $K^a_{\theta,h}$ and measurement law $M^a_h$. The counterfactual observable distribution is

$$F^a_{Y,\theta}(y \mid x) = \int M^a_h((-\infty, y] \mid x, x') K^a_{\theta,h}(\mathrm{d}x' \mid x).$$

This object is not identified by a good ECF fit alone. It additionally requires an explicit intervention rule, support or extrapolation restrictions, and an invariance statement saying which parts of the generator and measurement law are held fixed under the intervention. Without those assumptions, the paper reports predictive distributions and score gaps, not structural counterfactual distributions.

**Proposition 1** (Generator inversion as a distributional forecast target)**.** *Suppose a candidate generator $\theta$ and measurement law $M_h$ imply the observable distribution $F_{Y,\theta}(\cdot \mid x)$ in* (28)*. Then every forecast object used in the paper is a functional of this distribution: the conditional mean, conditional variance, quantiles, prediction intervals, tail probabilities, and binary rare-event probabilities. For*

*a negatively oriented proper scoring rule $S$, a benchmark forecast distribution $F_{B,i}$ fixed by a prespecified information rule, and a realized observable increment $y_i$, the population comparison target is*

$$\Delta_S(G, B) = \mathbb{E}[S(F_{B,i}, y_i) - S(F_{Y,\theta}(\cdot \mid X_i), y_i)].$$

*Positive values mean that the generator-implied observable distribution has lower expected score than the benchmark for the prespecified target, horizon, information set, and score. If the score gap is not positive, the generator may still fit the adjustment history, but it has not added observable distributional value for that forecast problem.*

*Proof.* The observable distribution determines its own moments, quantile function, interval endpoints, and event probabilities. Substituting those functionals into a proper scoring rule gives a forecast score for the generator-implied distribution. The displayed score gap is the expected benchmark score minus the expected generator score, so positive values favor the generator under a negatively oriented score. The final statement follows because adjustment-process fit and observable forecast value are distinct empirical claims. □

This comparison has a useful value condition. Let $\mathcal{I}_G$ be the information set used by the generator forecast after the adjustment state and measurement law are fixed, and let $\mathcal{I}_B \subseteq \mathcal{I}_G$ be the information set used by a reduced-form benchmark. Let

$$F_G^*(\cdot) = \Pr(Y_{t+h} - Y_t \leq \cdot \mid \mathcal{I}_G), \qquad F_B^*(\cdot) = \Pr(Y_{t+h} - Y_t \leq \cdot \mid \mathcal{I}_B)$$

denote the corresponding oracle predictive distributions for the observable increment.

**Proposition 2** (Forecast value requires incremental distributional information)**.** *For a strictly proper negatively oriented scoring rule $S$, the oracle forecast using $\mathcal{I}_G$ weakly improves expected score relative to the oracle forecast using $\mathcal{I}_B$:*

$$\mathbb{E}\{S(F_B^*, Y_{t+h} - Y_t)\} - \mathbb{E}\{S(F_G^*, Y_{t+h} - Y_t)\} \geq 0.$$

*The gap is zero if the benchmark information is sufficient for the observable distribution, meaning*

$$\Pr(Y_{t+h} - Y_t \leq y \mid \mathcal{I}_G) = \Pr(Y_{t+h} - Y_t \leq y \mid \mathcal{I}_B)$$

*for all relevant thresholds $y$. The oracle gap can be positive only if the adjustment information changes a distributional feature that the scoring rule can detect. A feasible empirical gap additionally requires the estimated generator to recover that feature accurately enough out of sample.*

*Proof.* Strict propriety means that, conditional on an information set, the true conditional distribution minimizes expected score. Since $\mathcal{I}_B \subseteq \mathcal{I}_G$, conditioning on $\mathcal{I}_G$ cannot increase the oracle expected score. If the two conditional distributions coincide, the two oracle forecasts are identical for the score and the gap is zero. If they differ on a feature to which the score is sensitive, the oracle gap is positive. With estimated forecasts the paper observes a feasible version of this comparison, so estimation error and regularization determine whether the oracle information advantage survives final-test evaluation. □

The proposition makes the paper's success criterion falsifiable. Generator inversion is not judged by whether the fitted object is more detailed than a reduced-form benchmark. It is judged by whether the estimated adjustment law improves or disciplines held-out observable distributions. Table 7 states the corresponding claim discipline. It separates adjustment-law fit, observable score improvements, uncertainty in score gaps, and the boundary between transition-law forecasts and generator claims. The same logic is used later when interpreting mixed characteristic-function and observable-score evidence.

The key econometric point is that aggregate equivalence does not imply generator equivalence. Two generators can imply the same drift and covariance for an observable while differing in event composition, skewness, tail risk, path probabilities, or counterfactual response to the state. A reduced-form mean or variance model can then fit the observed aggregate dynamics without identifying the conditional law that generates future distributional risk. Generator inversion targets the stronger object, but it must be evaluated through observable consequences.

**Proposition 3** (Aggregate equivalence does not imply generator equivalence)**.** *Suppose two event generators $\lambda$ and $\tilde{\lambda}$ have the same drift and covariance rates for an aggregate $Y = \chi(X)$ on a set of states. Let $\rho_r^Y(x) = \chi(x + \nu_r) - \chi(x)$ be the observable increment induced by event $r$. If there exists a frequency $u$ such that*

$$\sum_r \lambda_r(x)\{\exp(\mathrm{i}u^\top \rho_r^Y(x)) - 1\} \neq \sum_r \tilde{\lambda}_r(x)\{\exp(\mathrm{i}u^\top \rho_r^Y(x)) - 1\}$$

Table 7: Falsifiable claims in generator inversion.

The table states which claims are supported by different final-test evidence patterns. It separates observable distribution value from adjustment-law discipline, so an improved generator fit is not automatically reported as an improved observable forecast.

| Evidence on the final test block | Claim supported | Claim not supported |
|---|---|---|
| *Observable distribution value* | | |
| Generator-implied distributions improve proper observable scores relative to direct benchmarks, with a stable positive score gap. | The adjustment law contains predictive distributional information for the stated target, horizon, and information set. | The result does not show that the generator dominates for other targets, horizons, or domains. |
| Generator signals improve only a mean, volatility, quantile, interval, or tail projection. | The generator is useful for that forecast functional and can discipline the corresponding reduced-form projection. | The result is not evidence that the full observable distribution is better forecast. |
| Observable scores improve, but characteristic loss does not. | The generator-derived signal may improve the reported observable forecast through the measurement or projection law. | The result does not show that the adjustment law itself is better matched in the ECF sense. |
| *Adjustment-law discipline* | | |
| Characteristic loss improves, but observable scores do not. | The ECF restrictions identify adjustment-process information that may matter for other mappings or targets. | The result is not an observable forecasting gain for the reported outcome. |
| Characteristic loss improves only for selected frequencies or instruments. | The adjustment law matches those distributional features on the stated design. | The result is not a general improvement of the full conditional law unless other bands and instruments agree. |
| *Uncertainty and identification boundary* | | |
| The score gap is positive, but the bootstrap interval crosses zero or the positive share is weak. | The point estimate favors the generator and warrants reporting as suggestive evidence. | The result is not stable final-test evidence of generator value. |
| Only transition panels are observed and no semigroup or injective transition-to-generator restriction is imposed. | The fitted object supports an observed-horizon transition-law forecast. | The result should not be reported as identification of a continuous-time local generator. |
| Counterfactual distributions are reported with an explicit intervention rule, invariant components, and support restrictions. | The generator can be used as a disciplined counterfactual simulator for that stated regime. | The result does not support other interventions or structural counterfactuals outside the stated invariance and support conditions. |
| The anchor or a simple reduced-form benchmark wins validation or final testing. | The selected data and horizon do not justify the richer generator for that purpose. | The paper should not claim generator value beyond documenting where it fails. |

*on a set reached with positive probability, then the two generators are observationally equivalent for local Gaussian projections of $Y$, but they imply different conditional characteristic functions and can be separated by characteristic-function moments with informative instruments.*

*Proof.* The local Gaussian projection depends only on the first two cumulant rates. Equality of drift and covariance therefore makes the two generators equivalent for that projection. The conditional characteristic exponent depends on the full exponential transform of the event increments. If the two exponents differ at some frequency on a set with positive probability, the conditional characteristic functions differ over short intervals. Instruments that put positive weight on that set produce different population characteristic moments. □

This proposition explains why the empirical-characteristic-function criterion is natural. A likelihood or transition-fit benchmark asks whether the observed adjustments are probable under a candidate generator. The regularized ECF criterion asks whether the same generator reproduces conditional distributional features of aggregate increments across frequencies and state instruments. Validation then decides whether the ECF step improves held-out characteristic loss relative to the likelihood or transition anchor. The method therefore uses the generator as a disciplined distributional object rather than as an unrestricted latent state.

The inversion idea is related to the abduction logic in demand estimation. In static demand, Berry inversion recovers a latent demand index that makes market-specific counterfactuals possible (Berry, 1994; Berry and Haile, 2014; Pearl, 2009). Albrecht and Traina (2026) emphasize the gap between an average demand response and a market-specific demand curve. The dynamic version replaces the scalar latent index with a conditional adjustment generator. A direct forecast averages over the adjustment information in its information set, just as an average demand response averages over market-specific latent conditions. Generator inversion is the abduction step for dynamic systems: recover the state-specific adjustment law, then compute the observable distribution that would follow from that law. The object to be recovered is therefore not only a demand shifter or a fitted event model; it is the law that maps current states into future adjustment paths and, through those paths, into observable conditional distributions.

This broader interpretation covers several empirical environments. In an order book, the events are arrivals, executions, cancellations, and deletions. In a processor panel, the events may be price revisions, production reallocations, inventory changes, or benchmark-price transmissions. In credit,

labor, or firm dynamics, the events may be defaults, job transitions, investment adjustments, or entry and exit. The common feature is not the domain. It is the existence of event or transition data rich enough to estimate an adjustment law and test whether the law adds distributional information for observables.

## 3 Adjustment Generators and Observation Regimes

The general object is a conditional adjustment generator. Let $X_t$ denote the state observed before an adjustment and let $\mathcal{L}_\theta$ be an operator that maps test functions $f$ into local conditional changes,

$$\mathcal{L}_\theta f(x) = \lim_{\Delta \downarrow 0} \frac{\mathbb{E}_\theta[f(X_{t+\Delta}) \mid X_t = x] - f(x)}{\Delta}. \tag{2}$$

This definition covers event histories and coarser transition data. If every adjustment event is observed, $\mathcal{L}_\theta$ can be parameterized through event intensities. If only interval transitions are observed, a candidate generator is seen through the transition kernel $K_{\theta,\Delta}(\mathrm{d}x' \mid x)$ generated by the semigroup $\exp(\Delta\mathcal{L}_\theta)$. The econometric target is therefore not tied to a particular sampling frequency. It is the conditional law that maps the current state into future states, subject to the identification limits of the observation regime.

The observation regime determines which part of the generator is seen directly. With event histories, the data record the event type, event time, and pre-event state, so the compensator identifies local adjustment intensities on the exposed state support. With transition panels, the data record an initial state, a future state, and the elapsed horizon. The intervening events are latent, but the transition kernel still carries the generator's distributional implications. The ECF estimator uses this common object: it compares the realized increment with the conditional characteristic function implied by the candidate generator, whether that characteristic function is computed from event intensities, an analytic transition kernel, or simulated transition paths.

This distinction is important for identification. Transition panels directly identify transition laws at the horizons and states that are observed. Calling the object a generator estimate requires additional structure: observed event histories, multiple horizons tied together by a semigroup restriction, or a parameterization for which the map from generator to observed transition kernels is injective on the visited support. Without that structure, the estimand is the transition-law projection, and generator-based counterfactuals outside the observed horizons require an explicit modeling

assumption.

Table 8: Observation regimes for generator inversion.

The table compares event-history data and transition-panel data. In both cases, the estimator targets a conditional adjustment law, constructs a conditional characteristic function, and evaluates the observable distributions implied by the fitted generator.

| Observation regime | Data object | Generator anchor | Distributional restriction |
|---|---|---|---|
| Event histories | Event type, event time, pre-adjustment state, and exposure. | Counting-process likelihood for state-dependent adjustment intensities. | Conditional characteristic function computed from the fitted event intensities. |
| Transition panels | Initial state, future state, elapsed horizon, and observable outcome. | Transition likelihood, simulated transition fit, or prespecified transition benchmark. | Conditional characteristic function computed from the transition kernel or from simulated transition paths. |

For a continuous-time Markov jump system with state $n \in \mathbb{N}^d$, let $P(n,t) = \Pr(X_t = n)$ be the probability of state $n$ at time $t$. An event $r$ changes the state by a stoichiometric vector $\nu_r$ and occurs at intensity $\lambda_r(n)$. The master equation is

$$\partial_t P(n,t) = \sum_r \left[\lambda_r(n-\nu_r)P(n-\nu_r,t) - \lambda_r(n)P(n,t)\right]. \tag{3}$$

The generator is the organizing adjustment law, but its econometric value comes from the distributions it induces. It defines likelihoods for event histories, transition probabilities for coarser panels, moment equations, diffusion limits, and rare-event probabilities, all of which are useful only insofar as they identify, forecast, or test observable consequences.

Doi–Peliti notation rewrites this generator using creation and annihilation operators. For one component, $a^\dagger|n\rangle = |n+1\rangle$ and $a|n\rangle = n|n-1\rangle$. A transition that adds one unit contributes a factor $a^\dagger - 1$, while a transition that removes one unit contributes $1 - a^\dagger$ times the annihilation operator. For a probability state $|\Psi(t)\rangle = \sum_n P(n,t)|n\rangle$, the evolution can be written as

$$\frac{\mathrm{d}}{\mathrm{d}t}|\Psi(t)\rangle = \hat{L}|\Psi(t)\rangle. \tag{4}$$

### 3.1 Choice of Representation

The generator is invariant, but the representation is not innocuous. Different bases make different economic and statistical objects easy to see. The occupation-number basis is natural for exact event

likelihoods and finite-state calculations because the state variables are the observed counts. The generating-function basis turns the master equation into a partial differential equation and is useful for spectral methods, Wentzel–Kramers–Brillouin approximations, and rare-event calculations. The Poisson representation can be useful when count distributions are overdispersed or when one wants a continuous latent intensity representation. Coherent states expose the Hamiltonian and cumulant hierarchy. These alternatives are reviewed by Weber and Frey (2017), who emphasize that master equations can be represented through generating functions, Poisson representations, PDE mappings, spectral methods, rare-event asymptotics, and path integrals.

For econometrics this is a modeling choice, not only a mathematical convenience. If the data are event histories, the count basis is closest to the likelihood. If the target is conditional moments, the coherent-state Hamiltonian is compact. If the target is rare-event probabilities, an action representation is more direct. If the target is a low-dimensional state-space approximation, a Poisson or spectral basis may be preferable. The choice of basis therefore determines which approximation error is transparent. It also determines which quantities can be estimated without adding unnecessary latent structure. This is one reason why the paper treats Doi–Peliti as a language for moving between representations rather than as a single estimator. The broader stochastic-mechanics view of creation operators, coherent states, and stochastic Petri nets in Baez and Biamonte (2018) gives a useful mathematical reference point for this interpretation.

Passing to coherent states replaces the operator generator by a Hamiltonian $H(n, \varpi)$, where $\varpi$ is the conjugate response field. For a jump $\nu_r$, the contribution is

$$\lambda_r(n) \left( \exp(\varpi^\top \nu_r) - 1 \right). \tag{5}$$

Expanding around $\varpi = 0$ yields

$$H(n, \varpi) = \varpi^\top A(n) + \frac{1}{2} \varpi^\top B(n) \varpi + \sum_{k \geq 3} \frac{1}{k!} C_k(n)[\varpi^{\otimes k}], \tag{6}$$

where $A(n)$ is the drift, $B(n)$ is the conditional covariance rate, and the higher tensors $C_k(n)$ are conditional cumulant rates. Standard diffusion econometrics keeps the first two terms. The full jump generator keeps the hierarchy.

The same Hamiltonian defines a path action

$$S[n, \varpi] = \int_0^T \left[\varpi_t^\top \dot{n}_t - H(n_t, \varpi_t)\right] \mathrm{d}t. \tag{7}$$

After optimizing over $\varpi$, this gives the large-deviation cost of a path. This is where the Doi–Peliti language adds information beyond the local Gaussian approximation.

## 4 Characteristic Generators

The econometric object in this paper is not only the event intensity and not only a finite list of conditional moments. It is the conditional characteristic law implied by the event generator. For a jump system with adjustment types $r \in \mathcal{R}$, transition vectors $\nu_r$, and intensities $\lambda_{r,\theta}(x)$, define

$$H_\theta(x, \mathrm{i}u) = \sum_{r \in \mathcal{R}} \lambda_{r,\theta}(x)\{\exp(\mathrm{i}u^\top \nu_r) - 1\}. \tag{8}$$

Here $u$ has the same dimension as the state increment, i is the imaginary unit, and $H_\theta(x, \mathrm{i}u)$ is the coherent-state Hamiltonian evaluated at imaginary momentum. For a short interval $\Delta$, with the state held fixed over the interval,

$$\phi_\theta(u \mid x, \Delta) = \mathbb{E}_\theta[\exp\{\mathrm{i}u^\top (X_{t+\Delta} - X_t)\} \mid X_t = x] \approx \exp\{\Delta H_\theta(x, \mathrm{i}u)\}. \tag{9}$$

The approximation is exact for a compound Poisson step with frozen intensities and is the local characteristic function of the jump generator more generally.

For transition data the same object can be defined without observing the intervening events. Let $K_{\theta,\Delta}(\mathrm{d}x' \mid x)$ be the transition kernel implied by the generator over horizon $\Delta$. Then

$$\phi_\theta(u \mid x, \Delta) = \int \exp\{\mathrm{i}u^\top (x' - x)\} K_{\theta,\Delta}(\mathrm{d}x' \mid x). \tag{10}$$

The ECF estimator below only requires this conditional characteristic function. Event-history data supply it through the jump intensities; transition data supply it through a parametric or simulated transition law. This is what allows the same estimator to cover high-frequency events and lower-frequency panels.

Operationally, a transition-panel observation is a record $(x_i, x_i^+, \Delta_i)$. The initial state $x_i$ is observed

before the adjustment interval, $x_i^+$ is the state observed after the interval, and $\Delta X_i = x_i^+ - x_i$ is the realized state increment. If $K_{\theta,\Delta_i}$ is available in closed form, (10) gives $\phi_\theta(u \mid x_i, \Delta_i)$ directly. If the transition law is available only by forward simulation, the same moment condition uses simulated future states $X_{i,+}^{(s)}(\theta)$, $s = 1, \ldots, S_{\mathrm{tr}}$, and the Monte Carlo approximation

$$\hat{\phi}_{\theta,S_{\mathrm{tr}}}(u \mid x_i, \Delta_i) = \frac{1}{S_{\mathrm{tr}}} \sum_{s=1}^{S_{\mathrm{tr}}} \exp\{\mathrm{i}u^\top (X_{i,+}^{(s)}(\theta) - x_i)\}. \tag{11}$$

Here $S_{\mathrm{tr}}$ is the number of forward transition draws for each observed state and horizon. The estimation target is then the transition law at the observed states and horizons. Calling the fitted object a generator requires the additional semigroup or injectivity conditions stated in Section 3; otherwise the paper reports a transition-law forecast rather than a continuous-time generator claim outside the observed horizons.

This object gives the paper a sharper econometric interpretation. The likelihood or transition model identifies the adjustment law supported by the observation regime: event histories support a local generator, while transition panels first support the observed-horizon transition law unless the additional generator-identification conditions are imposed. Equation (8) then translates event rates, or more generally the transition law in (10), into a conditional characteristic function for aggregate changes. Differentiating $H_\theta(x, \mathrm{i}u)$ at $u = 0$ gives the conditional cumulants in the jump case. Keeping only the first two derivatives gives the Gaussian diffusion approximation. Evaluating the full exponential away from zero keeps jump asymmetry, discreteness, and tail information that a quadratic projection discards.

The characteristic-generator view connects the paper to empirical characteristic-function methods (Feuerverger and Mureika, 1977), continuum-moment GMM (Carrasco and Florens, 2000), and conditional-characteristic-function (CCF) estimation for dynamic models (Singleton, 2001). The distinction is the source of the conditional characteristic function. In affine diffusion applications, the CCF is often known from a parametric state process. Here it is induced by an estimated generator for the observed adjustment process. This lets the same economic object be evaluated through both likelihood and characteristic-function restrictions.

The estimator is built from $n$ observation triples $(x_i, \Delta X_i, \Delta_i)$, where $x_i$ is the pre-adjustment state, $\Delta_i$ is the elapsed horizon, and $\Delta X_i = X_{t_i+\Delta_i} - X_{t_i}$ is the realized state increment. Let $\mathcal{U} = \{u_1, \ldots, u_M\}$ be the chosen frequency grid, with $u_m$ conformable with $\Delta X_i$, and let $\mathcal{Z} =$

$\{z_0, \ldots, z_J\}$ be standardized state instruments with $z_0(x) \equiv 1$. For a complex number $c$, write $|c|^2 = \mathrm{Re}(c)^2 + \mathrm{Im}(c)^2$. The empirical moment component is

$$g_{mj,n}(\theta) = \frac{1}{n} \sum_{i=1}^{n} z_j(x_i) \left[ \exp\{\mathrm{i} u_m^\top \Delta X_i\} - \phi_\theta(u_m \mid x_i, \Delta_i) \right]. \tag{12}$$

The vector $g_n(\theta)$ stacks the real and imaginary parts of $g_{mj,n}(\theta)$ over all frequencies and standardized state instruments. The conditional empirical-characteristic-function (ECF) estimator starts from the minimum-distance target

$$\hat{\theta}_{\mathrm{ECF}} = \arg\min_{\theta} g_n(\theta)' W_n g_n(\theta). \tag{13}$$

Here $W_n$ is a positive semidefinite weighting matrix. Identity weighting gives the first-step estimator; a regularized inverse block long-run covariance matrix gives the second-step weighting rule. Table 9 collects the symbols that enter the criterion, including the sample dimensions, block length, moment dimension, validation block, and regularization objects.

To make the weighting rule reproducible, let $\psi_i(\theta)$ be the real-valued contribution after stacking the real and imaginary characteristic residuals over all selected frequencies and instruments, and let $\tilde{\psi}_i(\theta) = \psi_i(\theta) - n^{-1} \sum_{\ell=1}^{n} \psi_\ell(\theta)$. For block length $L$,

$$\hat{\Gamma}_0(\theta) = \frac{1}{n} \sum_{i=1}^{n} \tilde{\psi}_i(\theta) \tilde{\psi}_i(\theta)', \tag{14}$$

$$\hat{\Gamma}_\ell(\theta) = \frac{1}{n} \sum_{i=\ell+1}^{n} \tilde{\psi}_i(\theta) \tilde{\psi}_{i-\ell}(\theta)', \qquad \ell = 1, \ldots, L, \tag{15}$$

$$\hat{\Gamma}_n(\theta) = \hat{\Gamma}_0(\theta) + \sum_{\ell=1}^{L} \left(1 - \frac{\ell}{L+1}\right) \{\hat{\Gamma}_\ell(\theta) + \hat{\Gamma}_\ell(\theta)'\}. \tag{16}$$

The second-step weighting matrix is

$$W_n = \{\hat{\Gamma}_n(\hat{\theta}_\rho^{(1)}) + \kappa \bar{d} I\}^{-1}, \qquad \bar{d} = \mathrm{tr}\{\hat{\Gamma}_n(\hat{\theta}_\rho^{(1)})\}/p_\psi, \tag{17}$$

where $\hat{\theta}_\rho^{(1)}$ is the first-step fit, $p_\psi$ is the number of stacked real moments, and $\kappa > 0$ is a small numerical ridge. The matrix $I$ is the conformable identity matrix. The scalar $\bar{d}$ puts the ridge on the same scale as the average variance of the stacked moment contribution, so the inversion is a regularized weighting step rather than an unscaled numerical adjustment. The reported characteristic losses use the same moments on a held-out block. Let $\mathcal{T}$ be an evaluation block with $n_{\mathcal{T}}$ observations, let

Table 9: Notation for conditional ECF estimation.

The table collects the symbols used in the conditional ECF criterion, weighting rule, regularization step, adaptive design selection, and held-out characteristic-function loss.

| Symbol | Name | Calculation role |
|---|---|---|
| $n$, $M$, $J$ | Sample and grid dimensions | Number of estimation observations, number of frequencies, and number of non-intercept state instruments. |
| $(x_i, \Delta X_i, \Delta_i)$ | ECF observation | Pre-adjustment state, realized increment, and horizon for observation $i$. |
| $\mathcal{U}$, $u$ | Frequency grid and frequency | Arguments at which the conditional characteristic function is matched. |
| $\mathcal{Z}$, $z_j(x)$ | State instruments | Conditioning functions that make the characteristic restrictions state-dependent. |
| $\phi_\theta(u \mid x, \Delta)$ | Conditional characteristic function | Generator-implied transform of the state increment at frequency $u$, state $x$, and horizon $\Delta$. |
| $S_{\mathrm{tr}}$ | Simulated transition draws | Number of forward transition draws used to approximate $\phi_\theta$ when the transition law is available only by simulation. |
| $g_n(\theta)$ | Stacked characteristic moments | Sample average of instrumented differences between realized and generator-implied characteristic functions. |
| $\psi_i(\theta)$ | Moment contribution | Real-valued contribution after stacking real and imaginary characteristic residuals. |
| $L$, $p_\psi$, $I$ | Block length, moment dimension, and identity matrix | Dependence window for long-run covariance estimation, number of stacked real moments, and conformable identity matrix. |
| $W_n$ | Moment weighting matrix | Identity matrix in the first step; regularized inverse long-run covariance matrix in the second step. |
| $\hat{\Gamma}_n$, $\kappa$, $\bar{d}$ | Long-run covariance, ridge, and scale | Block covariance estimate for $\psi_i(\theta)$, numerical ridge, and average covariance scale used before inversion. |
| $\hat{\theta}_{\mathrm{A}}$ | Anchor estimate | Likelihood, transition-density, or prespecified fit used for initialization, comparison, and regularization. |
| $\rho$, $\Omega$, $p_\theta$ | Regularization weight, scaling, and parameter dimension | Validation-selected penalty, scaling matrix, and number of free generator coefficients used to control movement away from the anchor. |
| $\mathcal{P}$, $\hat{\rho}$ | Penalty grid and selected penalty | Finite set of candidate regularization weights and the value selected on the validation block. |
| $\mathcal{V}$, $n_{\mathcal{V}}$ | Validation block and size | Held-out block used to select the characteristic design and regularization before final testing. |
| $g_{\mathcal{V}}$, $W_{\mathcal{V}}$ | Validation moments and weighting | Same stacked characteristic moments and weighting rule, computed on the validation block. |
| $d \in \mathcal{D}$, $\xi$ | ECF design and complexity price | Candidate characteristic design and prespecified penalty for larger moment sets. |
| $\mathcal{U}_b$, $L_{\mathrm{CCF}}$ | Frequency band and characteristic-function loss | Held-out squared characteristic residuals averaged over instruments and the selected frequency band. |

$\phi_{a,i}(u)$ denote the conditional characteristic function from candidate model $a$ at observation $i$, and define

$$\eta_{a,i}(u) = \exp\{\mathrm{i}u^\top \Delta X_i\} - \phi_{a,i}(u).$$

For a frequency band $\mathcal{U}_b$, the conditional characteristic loss reported in the tables is the squared norm of the held-out instrumented characteristic residual:

$$L_{\mathrm{CCF}}(a,b) = \frac{1}{|\mathcal{U}_b|(J+1)} \sum_{u\in\mathcal{U}_b} \sum_{j=0}^{J} \left| \frac{1}{n_{\mathcal{T}}} \sum_{i\in\mathcal{T}} z_j(x_i)\eta_{a,i}(u) \right|^2. \tag{18}$$

The loss is nonnegative and lower values mean that the candidate model better matches the conditional characteristic function of held-out state increments at the selected frequencies and instruments. The unconditional characteristic loss is the same calculation with only the intercept instrument $z_0(x) \equiv 1$. In simulations, where the true generator is known, the oracle distance is

$$D_{\mathrm{oracle}}(a,b) = \frac{1}{|\mathcal{U}_b| n_{\mathcal{T}}} \sum_{u\in\mathcal{U}_b} \sum_{i\in\mathcal{T}} \left|\phi_{a,i}(u) - \phi_{0,i}(u)\right|^2, \tag{19}$$

where $\phi_{0,i}$ is the characteristic function implied by the data-generating generator.

The characteristic losses are therefore not only diagnostics. They define a first layer of testable restrictions: the estimated generator should reproduce the conditional characteristic law of held-out increments. The paper then asks a second question, namely whether the same fitted law improves the conditional distribution of observable outcomes relative to direct benchmarks. Table 10 summarizes this hierarchy and states how each reported measure should be interpreted.

In finite samples this target can be too free. Characteristic moments are noisy, especially once frequencies and instruments are stacked, and an unconstrained ECF step can move far away from an anchor that already captures the event or transition history. Let $\hat{\theta}_{\mathrm{A}}$ denote this anchor: a point-process likelihood fit when event histories are observed, a transition-likelihood fit when a transition density is available, or another prespecified transition fit when only coarser increments are observed. The implemented estimator therefore uses a validation-regularized version,

$$\hat{\theta}_\rho = \arg\min_\theta \left\{ g_n(\theta)' W_n g_n(\theta) + \rho(\theta - \hat{\theta}_{\mathrm{A}})' \Omega (\theta - \hat{\theta}_{\mathrm{A}}) \right\}, \tag{20}$$

Table 10: Testable restrictions and reported measures.

The table links the framework's testable questions to the restrictions being tested, the measures reported on validation or final-test data, and the interpretation of an improvement or failure. Panel headers separate characteristic-function tests, observable forecast tests, and projection or rare-event diagnostics.

| Empirical question | Restriction being tested | Reported measure | Interpretation |
|---|---|---|---|
| *Characteristic restrictions* | | | |
| Does the generator fit conditional increments? | Realized increments should have the same conditional characteristic function as the generator at the selected frequencies and states. | Held-out characteristic loss and bootstrap loss gap relative to the anchor or benchmark. | Improvement supports distributional fit for the adjustment law, not yet observable forecast value. |
| Should the ECF step move away from the anchor? | Additional frequencies, instruments, or weaker regularization must improve validation loss enough to offset sampling noise. | Selected characteristic design, selected penalty, validation loss, and final-test characteristic loss. | If the anchor wins, the richer ECF step is not reported as an empirical gain. |
| *Observable distribution restrictions* | | | |
| Does the generator improve observable forecasts? | The generator-implied observable distribution should score better than a prespecified direct benchmark for a fixed target and horizon. | Proper score gap, bootstrap interval, and share of positive block draws. | Positive gaps support value for that observable target; failures limit the claim to adjustment-law fit. |
| Does the generator improve distributional features? | Observable characteristic functions, quantiles, intervals, tail probabilities, or volatility implied by one generator distribution should improve the corresponding score. | Observable characteristic-function loss, pinball loss, interval score, tail-probability loss, ranking score, and volatility loss. | Evidence is feature-specific unless several scores improve under the same fixed design. |
| *Projection and rare-event restrictions* | | | |
| Do generator signals add value beyond aggregate controls? | Forecast equations augmented by generator-implied signals should improve held-out loss relative to aggregate projections. | Mean, volatility, quantile, or tail score gap against the direct projection. | Improvement shows predictive information in the adjustment layer; no improvement means the projection is sufficient for that target. |
| Are rare-state diagnostics credible? | Prespecified rare states should receive higher generator-implied risk scores than ordinary states. | Tail-probability score, area under the ranking curve, decile lift, and action diagnostic. | These are diagnostics for the stated rare event, not unrestricted stress-test optimality. |

where $\Omega$ is a positive diagonal scaling matrix. In the implementation this penalty is

$$\rho \frac{1}{p_\theta} \sum_{k=1}^{p_\theta} \left( \frac{\theta_k - \hat{\theta}_{\mathrm{A},k}}{s_k} \right)^2, \qquad s_k = \max\{|\hat{\theta}_{\mathrm{A},k}|, 1\}, \tag{21}$$

Here $p_\theta$ is the number of free generator coefficients in the fitted parameter vector. The display is equivalent to using $\Omega = p_\theta^{-1} \operatorname{diag}(s_1^{-2}, \ldots, s_{p_\theta}^{-2})$. The scaling makes the penalty comparable across coefficients whose anchor magnitudes differ. Let $\mathcal{P}$ denote the finite penalty grid; it includes $\rho = \infty$, interpreted as the anchor itself. For a separate validation block $\mathcal{V}$, define

$$\hat{V}(\rho) = g_{\mathcal{V}}(\hat{\theta}_\rho)' W_{\mathcal{V}} g_{\mathcal{V}}(\hat{\theta}_\rho), \qquad \hat{\rho} = \arg\min_{\rho \in \mathcal{P}} \hat{V}(\rho). \tag{22}$$

Here $g_{\mathcal{V}}$ is the same stacked characteristic-moment vector computed on the validation block, and $W_{\mathcal{V}}$ is the validation weighting matrix. The final estimator refits (20) on the estimation sample at $\hat{\rho}$. Thus the ECF step changes the generator only when held-out characteristic moments support the change. Proposition 11 states the corresponding finite-grid oracle property. If the anchor already gives the best validation loss, the procedure keeps it up to validation noise. If the ECF restrictions improve validation loss by a fixed gap, the procedure moves away from the anchor by the amount selected out of sample.

The same validation logic also makes the ECF design adaptive without turning the final test into a specification search. Let $d \in \mathcal{D}$ index a finite set of prespecified characteristic designs. A design $d$ contains a frequency grid $\mathcal{U}_d$, an instrument set $\mathcal{Z}_d$, a block length, and a weighting rule. For each pair $(d, \rho)$, let $\hat{\theta}_{d,\rho}$ be the corresponding regularized ECF fit and let $p_{\psi,d}$ be the number of stacked real moments. The adaptive validation score is

$$\hat{V}(d, \rho) = g_{\mathcal{V},d}(\hat{\theta}_{d,\rho})' W_{\mathcal{V},d} g_{\mathcal{V},d}(\hat{\theta}_{d,\rho}) + \xi \frac{p_{\psi,d}}{n_{\mathcal{V}}}, \qquad (\hat{d}, \hat{\rho}) = \arg\min_{d \in \mathcal{D}, \rho \in \mathcal{P}} \hat{V}(d, \rho). \tag{23}$$

The constant $\xi \geq 0$ is a prespecified complexity price. With $\xi = 0$, selection is pure validation. With $\xi > 0$, the rule favors smaller moment sets unless additional frequencies and instruments improve held-out characteristic loss enough to pay for their sampling noise. In either case, the final test block is used only after $(\hat{d}, \hat{\rho})$ has been selected.

Table 11 states the implemented estimator as an algorithm. The key discipline is that the candidate frequency grids, instruments, block lengths, weighting rules, and regularization grid are fixed before

the final test block is evaluated. The ECF step is therefore not an unrestricted search for a better-looking generator. It is a validation-selected movement away from an interpretable anchor, with the option to keep a low-dimensional ECF design when high-frequency or heavily instrumented moments are too noisy.

Table 11: Adaptive regularized conditional empirical-characteristic-function estimation algorithm. The table lays out the estimator from data splitting to anchor fitting, characteristic-moment construction, two-step weighting, validation selection, final refitting, and final test reporting. Panel headers separate objects fixed before final testing, the anchor and ECF fit, and validation/reporting.

| Stage | Operation | Output |
|---|---|---|
| *Objects fixed before final testing* | | |
| Data split | Partition the ordered observations into estimation, validation, and final test blocks. | Separate samples for fitting, regularization choice, and final evidence. |
| Design library | Prespecify candidate frequency bands, standardized state instruments, block lengths, weighting rules, and regularization grid. | A finite library of characteristic designs and penalties that can be selected only on validation data. |
| *Anchor and ECF fit* | | |
| Anchor fit | Estimate the likelihood, transition-density, or prespecified transition benchmark on the estimation block. | An anchor generator used for initialization, comparison, and regularization. |
| Characteristic moments | Compute instrumented residuals between realized increments and generator-implied characteristic functions. | Stacked real and imaginary moments for the selected frequencies and instruments. |
| First-step ECF | Minimize the regularized ECF criterion with identity weighting for each design and penalty value. | Candidate generators indexed by characteristic design and regularization weight. |
| Second-step ECF | Estimate a block long-run covariance matrix for each stacked moment design and repeat the criterion with covariance weighting. | Dependence-adjusted candidate generators for the same design and penalty grid. |
| *Validation and reporting* | | |
| Validation selection | Evaluate candidate generators on the validation block using held-out characteristic loss plus any prespecified complexity price. | Selected characteristic design and regularization weight, including the possibility of keeping the anchor. |
| Final refit | Refit the selected regularized ECF estimator on the estimation block under the selected design and penalty. | Final generator used to construct observable predictive distributions. |
| Final test | Score the selected generator and reduced-form benchmarks on the untouched test block. | Characteristic losses, observable distribution scores, rare-event scores, and benchmark gaps. |

The three sample blocks have different evidentiary roles. The estimation block can produce a fitted adjustment law, but by itself it cannot show that the ECF step adds distributional value. The validation block can choose among prespecified characteristic designs and regularization values, but it cannot be used as final evidence after that choice has been made. The final test block is the only block used to support the paper's reported characteristic-loss gaps, observable forecast-score gaps, and benchmark conclusions. This separation is what turns adaptive ECF fitting into a controlled estimator rather than a specification search over frequencies, instruments, and scores.

The first step uses identity weighting. The second step applies (16)–(17) conditional on the selected regularization. This is a finite-dimensional version of conditional-characteristic-function GMM. The

anchor remains useful: it initializes the generator coefficients, supplies a benchmark, anchors the regularized ECF estimator, and provides a comparison target for held-out characteristic loss.

A model that captures the conditional distribution should make these moments small across frequencies and conditioning functions. Low frequencies emphasize mean and variance restrictions. Intermediate frequencies reveal skewness, discreteness, and higher cumulants. Very high frequencies can be noisy in finite samples or uninformative when all candidate characteristic functions are close to zero. The empirical criterion is therefore not that the full generator representation must dominate at every frequency. The criterion is whether adjustment-generator information reduces characteristic-function loss where aggregate projections omit economically relevant state information.

# 5 From Generators to Observable Outcomes

The generator is not the final empirical object. Economists usually want predictions and counterfactuals for observables: prices, quantities, inventories, market shares, pass-through, or stress states. The role of the generator is to supply structured conditional signals for those observables. It is useful only if the estimated adjustment law improves the measurement, prediction, or counterfactual analysis of outcomes the researcher can observe.

Let $Y_t$ denote an observable outcome, such as a transaction price, a processor payout price, a delivered quantity, an inventory stock, an employment state, or a default indicator. Each adjustment event $r \in \mathcal{R}$ changes the latent state and may also change $Y_t$ directly. If $\rho_r^Y(x)$ is the event-specific change in the observable at state $x$, the generator implies local observable projections

$$A_Y(x) = \sum_{r \in \mathcal{R}} \rho_r^Y(x) \lambda_r(x), \tag{24}$$

$$B_Y(x) = \sum_{r \in \mathcal{R}} \rho_r^Y(x) \rho_r^Y(x)^\top \lambda_r(x). \tag{25}$$

Here $A_Y(x)$ is the local conditional drift rate of the observable and $B_Y(x)$ is the local second-moment rate. When $Y_t$ is scalar, both objects are scalar except that $B_Y(x)$ is nonnegative; when $Y_t$ is vector-valued, $A_Y(x)$ is a vector and $B_Y(x)$ is the corresponding second-moment matrix. These are not new observables; they are generator-implied pressure and risk measures constructed from

the observed adjustment history. A reduced-form model can then be written for the actual outcome,

$$Y_{t+h} - Y_t = \alpha_h + \beta_h^\top A_Y(x_t) + \gamma_h^\top B_Y(x_t) + \delta_h^\top w_t + \varepsilon_{t+h}, \tag{26}$$

Here $\alpha_h$ is the horizon-specific intercept, $\beta_h$, $\gamma_h$, and $\delta_h$ are forecast loadings, $w_t$ contains aggregate controls, and $\varepsilon_{t+h}$ is the remaining forecast error or measurement shock. The dimensions of the loading vectors conform to the dimensions of $A_Y(x_t)$, $B_Y(x_t)$, and $w_t$. In this equation the arrival and removal rates are not the object of economic interest. They are inputs into observable price, quantity, inventory, or stress predictions.

This bridge is also the clean comparison with vector autoregressions, error-correction models, and volatility models. A vector autoregression projects $Y_{t+h} - Y_t$ directly on lagged observables. The generator projection augments or restricts that regression with adjustment signals implied by the event law. If those signals do not improve observable loss on held-out data, the event layer has not added empirical value for that target. If they improve price changes, stress probabilities, or tail concentration, the generator has identified information that an aggregate projection missed.

The same object can be used for the full conditional distribution of an observable implied by the selected generator and measurement law. The target is

$$F_{Y,t+h}(y \mid x_t) = \Pr(Y_{t+h} - Y_t \leq y \mid x_t). \tag{27}$$

Let $M_h(\mathrm{d}y \mid x, x')$ denote the observable measurement or projection law over horizon $h$, conditional on the starting state $x$ and the future adjustment state $x'$. If the observable is a deterministic function of the state, $M_h$ is a point mass. If the generator supplies adjustment signals that enter an outcome equation, $M_h$ includes the residual uncertainty in that projection. With transition kernel $K_{\theta,h}$, the generator-implied observable distribution is

$$F_{Y,\theta}(y \mid x) = \int M_h((-\infty, y] \mid x, x') K_{\theta,h}(\mathrm{d}x' \mid x). \tag{28}$$

Equation (28) is the bridge from generator estimation to observable distributional forecasting. The local objects $A_Y(x)$ and $B_Y(x)$ are only its first two projections. The empirical content of the distribution is therefore tied to the information in the estimated generator and to the credibility of the measurement law; forecast scores test that joint object on held-out observations.

The measurement or projection law is therefore part of the empirical design, not a harmless afterthought. Table 12 separates the main cases. If the observable is a deterministic function of the state, held-out forecast errors mostly test the transition law. If an outcome equation or residual distribution is added, forecast scores test the generator and that observable mapping jointly. This distinction matters for interpretation: a generator can fit the adjustment process and still fail as an observable forecasting device if the mapping from adjustment states to outcomes is weak or misspecified.

Table 12: Observable mappings from generators to forecast targets.

The table distinguishes direct mappings, projection mappings, and limited or misspecified mappings from the estimated generator to observable forecast distributions. It clarifies which part of the framework is tested when observable scores improve or deteriorate.

| Observable mapping | Forecast construction | Interpretation of forecast evidence |
|---|---|---|
| *Direct mappings* | | |
| Deterministic state outcome | The future observable is a known function of the future adjustment state. | Forecast errors primarily test the estimated transition law of the adjustment state. |
| Event-level observable increments | Each adjustment event has a specified effect on the observable outcome. | Forecast scores test both event intensities and the event-to-outcome increment specification. |
| *Projection mappings* | | |
| Outcome equation with generator signals | Generator-implied drift, variance, quantiles, or simulated paths enter an observable forecast equation. | Forecast gains measure whether generator signals add predictive information beyond aggregate controls. |
| Residual distribution added to generator paths | Forward generator simulations are combined with residual uncertainty from the outcome projection. | Distributional scores test the generator, the projection equation, and the residual law jointly. |
| *Limited or misspecified mappings* | | |
| Latent or incomplete measurement | The mapping from adjustment state to observable outcome is only partly observed or must be approximated. | Weak observable scores may reflect measurement error rather than a useless adjustment generator. |
| Direct observable benchmark dominates | Reduced-form forecasts score better than the generator-implied observable distribution. | The relevant observable information is already captured by direct histories, or the generator-to-observable mapping is misspecified. |

Table 13 spells out the construction used in applications. The important point is that the same fitted generator produces all forecast summaries. When the observable mapping is estimated, for example through an outcome equation and residual distribution, that mapping is estimated on the training block before final testing. The mean forecast, volatility forecast, quantiles, intervals, and tail probabilities are not separate models unless the researcher deliberately replaces the generator distribution with a reduced-form benchmark.

Once the generator and the observable mapping are specified, $F_{Y,\theta}$ can be computed analytically in simple cases or simulated by drawing forward paths from the estimated generator. For held-out observation $i$, let $h_i$ be the forecast horizon and draw $S_{\text{sim}}$ future states $X_{i,h}^{(s)}$ from $K_{\hat{\theta},h_i}(\cdot \mid x_i)$.

Table 13: Constructing observable forecast distributions from a fitted generator.

The table shows how an estimated generator is translated into a predictive distribution for an observable outcome, and how that distribution is then reduced to forecast summaries and scored against direct reduced-form benchmarks.

| Step | Object constructed | Purpose |
|---|---|---|
| 1 | Starting state and horizon | Fix the pre-adjustment state $x_i$, forecast horizon $h_i$, and observable target $Y_{t_i+h_i} - Y_{t_i}$. |
| 2 | Adjustment transition law | Use the estimated event generator or transition kernel to obtain $K_{\hat{\theta},h_i}(\mathrm{d}x' \mid x_i)$. |
| 3 | Observable measurement law | Specify or estimate $M_{h_i}(\mathrm{d}y \mid x_i, x')$, either as a deterministic state mapping or as a projection law with residual uncertainty. |
| 4 | Predictive distribution | Combine $K_{\hat{\theta},h_i}$ and $M_{h_i}$ through (28), analytically or by forward simulation. |
| 5 | Forecast summaries | Extract conditional means, variances, quantiles, prediction intervals, and tail probabilities from the same distribution. |
| 6 | Benchmark comparison | Score the generator distribution against direct observable models using the untouched test block. |

Draw the corresponding observable increment $\tilde{y}_i^{(s)}$ from $M_{h_i}(\cdot \mid x_i, X_{i,h}^{(s)})$, and form

$$\widehat{F}_{G,i}(y) = \frac{1}{S_{\mathrm{sim}}} \sum_{s=1}^{S_{\mathrm{sim}}} \mathbb{1}\{\tilde{y}_i^{(s)} \leq y\}. \tag{29}$$

This empirical predictive distribution is the object from which the reported generator quantiles, prediction intervals, tail probabilities, and distribution scores are extracted. This makes the comparison with quantile regression direct. Quantile regression estimates selected conditional quantiles of the observable without modeling the adjustment law (Koenker and Bassett, 1978). The generator instead estimates the adjustment law first and then derives all quantiles from one coherent predictive distribution.

The empirical comparison should therefore use distributional forecast scores rather than only mean-squared errors. For a quantile level $\tau$, the pinball loss evaluates the forecasted quantile directly. Prediction intervals can be compared by their coverage, width, and interval score. Tail probabilities can be evaluated with probability losses, and a quantile-grid approximation to the continuous ranked probability score (CRPS) summarizes the full distribution (Gneiting and Raftery, 2007). These scores keep the paper's discipline intact: the generator matters only where it improves observable distribution forecasts relative to reduced-form distributional benchmarks.

Table 14 collects the notation used in the forecast-score calculations. The table separates observable

forecast distributions, realized outcomes, observable characteristic functions, quantile and interval summaries, rare-event probabilities, point forecasts, benchmark score gaps, and bootstrap quantities. The formulas below then define each measure using this notation.

All reported forecast measures follow the same calculation discipline. The generator, observable mapping, ECF design, regularization, benchmark library, benchmark information set, and benchmark selection rule are fixed before the final test block is scored. Each score is an average over the same held-out observations for a prespecified target and horizon. Losses are negatively oriented, so lower values are better; ranking scores are positively oriented, so higher values are better. Benchmark gaps are reported with a common sign convention: positive values favor the generator-implied observable distribution for the stated target, information set, horizon, and score.

Table 15 summarizes the evaluation measures used in the framework before the simulation reports the subset available in the diagnostic exercise. The table is not an additional estimator. It is a reader-facing map from the forecast object to the calculation and to the direction in which the measure is interpreted. The result tables use the same names and sign conventions. When a table rescales a loss for readability, the caption or float description states the scale; otherwise the displayed entry is the sample average defined in this section.

The simulation section also reports first-stage diagnostics, economic projection errors, rare-action diagnostics, and resampling summaries. Table 16 collects those calculations separately. This prevents the main forecast-score table from mixing standard scoring rules with simulation-only recovery checks.

The diagnostic table uses a separate simulation convention. Hats denote fitted quantities from candidate model $m$, and subscript 0 denotes the simulated truth when the truth is available. In projection formulas, $G$ is a placeholder for the generator-implied object being checked, such as an excess-demand drift, variance, or cumulant rate. The action symbols $\hat{I}_{m,i}$ and $I_{0,i}$ refer only to the prespecified fixed path used in the rare-event diagnostic. Bootstrap draws are indexed by $b = 1, \ldots, B_{\text{boot}}$, and full Monte Carlo replications are indexed by $j = 1, \ldots, N_{\text{rep}}$. Losses $L_{G,j}$ and $L_{B,j}$ are generator and benchmark losses in replication $j$; positively oriented scores $S_{G,j}$ and $S_{B,j}$ use the same convention.

Table 17 gives the interpretation rule for mixed evidence. This is important because the framework deliberately evaluates more than one object. A generator may improve the adjustment distribution

Table 14: Notation for observable forecast evaluation.

The table introduces the symbols used to score generator-implied and reduced-form forecast distributions for observable outcomes.

| Symbol | Name | Calculation role |
|---|---|---|
| $F_{Y,\theta}(\cdot \mid x)$ | Generator-implied observable distribution | Conditional distribution of the observable increment induced by the estimated adjustment law. |
| $\widehat{F}_{G,i}$ | Simulated generator forecast distribution | Empirical predictive distribution in (29), constructed from forward generator paths for held-out observation $i$. |
| $F_{m,i}$ | Model forecast distribution | Predictive distribution from model $m$ for held-out observation $i$. |
| $\varphi^{Y}_{m,i}(v)$ | Observable characteristic function | Characteristic function of $F_{m,i}$ at observable frequency $v$. |
| $\mathcal{U}_Y$, $\omega_\ell$ | Observable frequency grid and test functions | Frequencies and standardized conditioning functions used in observable characteristic-function loss. |
| $L_{\mathrm{OCF}}(m)$ | Observable characteristic-function loss | Squared instrumented characteristic residuals for the realized observable increment and the forecast distribution. |
| $m$, $G$, $B$, $S$ | Model and score labels | Generic model index, generator forecast, benchmark forecast, and scoring rule used for the reported comparison. |
| $\mathcal{A}_{\mathrm{dec}}$, $a^*_G$, $a^*_B$, $\ell(a,y)$, $\Delta_\ell(G,B)$ | Decision notation | Optional action set, forecast-implied generator and benchmark decisions, realized decision loss, and benchmark-minus-generator decision-value gap. |
| $S_{\mathrm{sim}}$, $\tilde{y}_i^{(s)}$ | Simulation size and draw | Number of forward generator paths and the $s$-th simulated observable increment for held-out observation $i$. |
| $h_i$ | Forecast horizon | Horizon over which the held-out observable increment and generator forecast are evaluated. |
| $\mathcal{T}$, $n_{\mathcal{T}}$ | Test block and size | Held-out observations used to compute characteristic losses, forecast scores, and benchmark gaps. |
| $y_i$ | Realized observable increment | Observed value against which the forecast distribution is scored. |
| $\mathcal{A}$ | Quantile grid | Ordered probability levels used for pinball loss and the quantile-grid distribution score. |
| $q_{m,i}(\tau)$ | Forecasted quantile | $\tau$-quantile of the predictive distribution from model $m$. |
| $L_\tau$, $\bar{L}_m$ | Pinball loss and average pinball loss | Quantile loss at level $\tau$ and its average over the test block and quantile grid. |
| $C_m$ | Quantile crossing share | Share of forecasts whose adjacent reported quantiles are not ordered monotonically. |
| $[\ell_{m,i,\alpha}, u_{m,i,\alpha}]$ | Prediction interval | Central $1-\alpha$ interval used for coverage, width, and interval-score diagnostics. |
| $S_{\alpha,i}(m)$ | Interval score | Proper score combining interval width and penalties for misses outside the interval. |
| $c$, $L_{\mathrm{tail}}(m;c)$ | Tail threshold and probability loss | Training-block lower-tail cutoff and squared loss for the predicted tail probability. |
| $R_i$, $\pi_{m,i}$ | Rare-event outcome and probability | Realized rare-state indicator and predicted probability used in probability and ranking scores. |
| $N_1$, $N_0$ | Rare and non-rare counts | Numbers of rare-state and non-rare observations in the final test block. |
| $\mathrm{AUC}_m$ | Area-under-curve score | Ranking probability that a rare-state observation receives a higher predicted risk than a non-rare observation. |
| $\mathcal{D}_m$ | Highest-risk decile | Held-out observations with the largest predicted rare-event probabilities under model $m$. |
| $\bar{R}$, $\bar{R}_{\mathcal{D}_m}$, $N_{\mathcal{D}_m}$ | Rare-state rates and count | Full test-block rare-state rate, highest-risk decile rare-state rate, and highest-risk decile rare-state count. |
| $\mathrm{Lift}_m$ | Decile lift | Rare-state rate in the highest-risk decile divided by the full test-block rare-state rate. |
| $\hat{y}_{m,i}$ | Point forecast | Scalar forecast used when a reported target is evaluated by root mean squared error. |
| $\mathrm{RMSE}_m$ | Root mean squared error | Square root of the average squared point-forecast error for model $m$. |
| $\Delta_S(G,B)$, $\hat{\Delta}_S(G,B)$ | Benchmark score gap | Population and sample differences between benchmark and generator scores under scoring rule $S$; positive values favor the generator. |
| $B_{\mathrm{boot}}$, $\hat{\Delta}_S^{*(b)}$ | Bootstrap draws and gaps | Number of block-bootstrap draws and the score gap recomputed in bootstrap draw $b$. |

Table 15: Calculation summary for evaluation measures.

The table lists the characteristic-function, observable-distribution, rare-event, and point-forecast measures used in the paper. Panel headers group related measures, and the final column states whether lower, higher, positive, or near-nominal values are preferred.

| Evaluation measure | Calculation | Object evaluated | Better value |
|---|---|---|---|
| *Characteristic-function diagnostics* | | | |
| Conditional characteristic-function loss | $L_{\mathrm{CCF}}$ in (18): squared instrumented characteristic residuals, averaged across a frequency band. | Conditional distribution of state increments. | Lower |
| Unconditional characteristic-function loss | Same calculation as $L_{\mathrm{CCF}}$, using only the intercept instrument. | Marginal distribution of state increments. | Lower |
| Observable characteristic-function loss | $L_{\mathrm{OCF}}$ in (31): squared instrumented residuals between the realized observable transform and the forecast transform. | Conditional distribution of observable increments. | Lower |
| Oracle distance | $D_{\mathrm{oracle}}$ in (19): mean squared distance from the true characteristic function. | Simulation-only distance from the data-generating law. | Lower |
| *Observable distribution forecasts* | | | |
| Average pinball loss | $\bar{L}_m$ in (33): average quantile loss over the reported quantile grid. | Forecasted conditional quantiles of the observable. | Lower |
| Continuous ranked probability score approximation | $2\bar{L}_m$, the equal-weight quantile-grid approximation. | Full predictive distribution of the observable. | Lower |
| Interval score | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} S_{\alpha,i}(m)$, using (34). | Width and calibration of a central prediction interval. | Lower |
| Coverage | Share of test observations inside the reported prediction interval. | Interval calibration at the stated nominal level. | Near nominal |
| Average width | Mean upper endpoint minus lower endpoint of the prediction interval. | Sharpness of the interval forecast, conditional on calibration. | Lower |
| Tail probability loss | $L_{\mathrm{tail}}(m; c)$ in (35). | Forecasted probability of a lower-tail observable outcome. | Lower |
| Quantile crossing share | $C_m$: share of forecasts with decreasing adjacent quantiles. | Internal coherence of reported quantiles. | Lower |
| Simulation paths | Number of forward generator paths used to approximate $F_{m,i}$. | Monte Carlo precision of the generator distribution forecast. | Not a score |
| *Rare-event and point-forecast diagnostics* | | | |
| Brier score | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (R_i - \pi_{m,i})^2$. | Rare-state probability forecast. | Lower |
| Area under receiver-operating-characteristic curve | $\mathrm{AUC}_m$ in (36): probability that a rare-state observation is ranked above a non-rare observation, with ties counted as one half. | Rare-state ranking. | Higher |
| Rare-state rate | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} R_i$. | Held-out prevalence of the rare event. | Not a score |
| Highest-risk decile rate | $\lvert\mathcal{D}_m\rvert^{-1} \sum_{i \in \mathcal{D}_m} R_i$. | Rare-state concentration among the observations ranked as highest risk. | Higher |
| Highest-risk decile lift | $\mathrm{Lift}_m$: rare-state rate in the highest-risk decile divided by the full test-sample rare-state rate. | Tail concentration of stress forecasts. | Higher |
| Root mean squared error | $\mathrm{RMSE}_m$: square root of mean squared point-forecast error. | Conditional mean or second-moment projection. | Lower |
| Benchmark gap | Benchmark loss minus generator loss; for ranking, generator score minus benchmark score. | Relative value of generator information. | Positive |

Table 16: Calculation summary for simulation diagnostics.

The table defines the event-generator, projection, rare-action, bootstrap, and Monte Carlo diagnostics used in the simulation section. Panel headers separate first-stage fit, generator projections, and uncertainty or rare-event diagnostics.

| Reported diagnostic | Calculation | Interpretation |
|---|---|---|
| In this table, $m$ indexes a fitted candidate, $\mathcal{T}$ is the held-out block, and $n_{\mathcal{T}}$ is its size. Hats denote fitted quantities. Subscript 0 denotes simulated truth, $G$ is the generator-implied projection being checked, $B_{\text{boot}}$ is the number of bootstrap draws, and $N_{\text{rep}}$ is the number of Monte Carlo replications. | | |
| *Event-generator fit* | | |
| Held-out event log likelihood | $n_{\mathcal{T}}^{-1} \sum_{r \in \mathcal{R}} \ell_{r,\mathcal{T}}(\hat{\theta}_{r,m})$, using the point-process log likelihood in (49). | Higher values indicate better event prediction. |
| True-intensity error | $\left[(n_{\mathcal{T}}\lvert\mathcal{R}\rvert)^{-1} \sum_{i \in \mathcal{T}} \sum_{r \in \mathcal{R}} \{\hat{\lambda}_{r,m}(x_i) - \lambda_{r,0}(x_i)\}^2\right]^{1/2}$. | Simulation-only recovery check; lower values are better. |
| Standardized martingale residual | $\max_r \lvert\hat{Z}_{r,m}\rvert$, where $\hat{Z}_{r,m} = \hat{M}_{r,m}/\sqrt{\hat{V}_{r,m}}$, $\hat{M}_{r,m} = \sum_{i \in \mathcal{T}} \{\mathbb{1}\{r_i = r\} - \hat{\lambda}_{r,m}(x_i)\Delta_i\}$, and $\hat{V}_{r,m} = \sum_{i \in \mathcal{T}} \hat{\lambda}_{r,m}(x_i)\Delta_i$. | Values closer to zero indicate better compensator fit. |
| Residual root mean square | $\left(\lvert\mathcal{R}\rvert^{-1} \sum_{r \in \mathcal{R}} \hat{Z}_{r,m}^2\right)^{1/2}$. | Summarizes residual size across event types. |
| Residual autocorrelation | $\max_r \lvert\hat{a}_{r,m}\rvert$, where $\hat{a}_{r,m}$ is the lag-one autocorrelation of $\hat{m}_{r,i,m} = \mathbb{1}\{r_i = r\} - \hat{\lambda}_{r,m}(x_i)\Delta_i$ on the held-out block. | Values closer to zero indicate less remaining serial structure. |
| *Generator projections* | | |
| Drift and variance error | $\left[n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\hat{G}_m(x_i) - G_0(x_i)\}^2\right]^{1/2}$, with $G = A_E$ for drift and $G = B_E$ for variance. | Simulation-only projection recovery check; lower values are better. |
| Projection forecast error | $\left[n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (y_i - \hat{y}_{m,i})^2\right]^{1/2}$, using the target-specific realized increment and forecast. | Held-out observable or moment forecast loss. |
| Cumulant increment error | $\left[n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\hat{C}_{k,m}^E(x_i)\Delta_i - C_{k,0}^E(x_i)\Delta_i\}^2\right]^{1/2}$, for cumulant order $k$. | Checks whether non-Gaussian cumulants are recovered. |
| Standardized skewness and excess kurtosis | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \lvert C_3^E(x_i)\Delta_i / \{C_2^E(x_i)\Delta_i\}^{3/2}\rvert$ and the held-out median of $C_4^E(x_i)\Delta_i / \{C_2^E(x_i)\Delta_i\}^2$. | Scale-free non-Gaussianity diagnostics. |
| Price-response slope error | $\left[n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\partial_p \hat{A}_{E,m}(x_i) - \partial_p A_{E,0}(x_i)\}^2\right]^{1/2}$. | Lower values indicate better recovery of the economic response. |
| Price-response bias | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\partial_p \hat{A}_{E,m}(x_i) - \partial_p A_{E,0}(x_i)\}$. | Shows whether the response magnitude is overstated or understated. |
| Sign agreement | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \mathbb{1}\{\text{sign}(\partial_p \hat{A}_{E,m}(x_i)) = \text{sign}(\partial_p A_{E,0}(x_i))\}$. | Higher values indicate better qualitative recovery. |
| *Rare-event and uncertainty diagnostics* | | |
| Action error | $\left[n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (\hat{I}_{m,i} - I_{0,i})^2\right]^{1/2}$, where $\hat{I}_{m,i}$ is the fitted fixed-path action and $I_{0,i}$ is the true fixed-path jump action. | Simulation-only rare-event geometry check; lower values are better. |
| Action ratio | $\text{median}_{i \in \mathcal{T}}(\hat{I}_{m,i}/I_{0,i})$. | Values close to one indicate correct action scale. |
| Rank correlation | $\text{Corr}\{\text{rank}(\hat{I}_{m,i}), \text{rank}(I_{0,i})\}$. | Higher values indicate better stress-state ranking. |
| Bootstrap interval | $[q_{0.025}\{\hat{\Delta}_S^{*(b)}\}, q_{0.975}\{\hat{\Delta}_S^{*(b)}\}]$, where $q_\alpha$ is the empirical $\alpha$-quantile across block-bootstrap score gaps. | Reports uncertainty for held-out benchmark gaps. |
| Positive bootstrap share | $B_{\text{boot}}^{-1} \sum_{b=1}^{B_{\text{boot}}} \mathbb{1}\{\hat{\Delta}_S^{*(b)} > 0\}$, where $B_{\text{boot}}$ is the number of bootstrap draws. | Higher values indicate more stable generator gains. |
| Monte Carlo ratio and better share | $\text{median}_j(L_{G,j}/L_{B,j})$ for loss ratios; $N_{\text{rep}}^{-1} \sum_{j=1}^{N_{\text{rep}}} \mathbb{1}\{L_{G,j} < L_{B,j}\}$ for better shares, where $L_{G,j}$ and $L_{B,j}$ are generator and benchmark losses. For likelihood and ranking scores, replace the ratio by $S_{G,j} - S_{B,j}$. | Ratios below one, differences above zero, and larger better shares favor the generator. |

without improving the observable forecast, or it may improve a tail forecast without lowering average loss. Those cases are not failures of the framework. They are evidence about where the adjustment law carries information and where the direct observable benchmark is already sufficient.

Table 17: Interpreting generator evidence against reduced-form benchmarks.

The table explains how to read combinations of characteristic-function diagnostics and observable forecast scores. It separates evidence that the adjustment law is better matched from evidence that the induced observable distribution is useful for forecasting.

| Evidence pattern | Interpretation | Reporting implication |
|---|---|---|
| *Joint characteristic and observable evidence* | | |
| Characteristic loss and observable scores improve | The adjustment law contains distributional information that survives the mapping to the observable outcome. | Treat the generator as empirically useful for the stated target, horizon, and information set. |
| Characteristic loss improves, but observable scores do not | The generator fits the adjustment process better, but the improved adjustment distribution is not yet useful for the reported observable. | Report the result as adjustment-process information, not as evidence of better observable forecasting. |
| Observable scores improve, but characteristic loss does not | The observable mapping or forecast projection extracts useful signals even though the full adjustment law is not better matched in the selected frequency grid. | Interpret the gain as forecast value from generator signals; check the frequency grid, instruments, and measurement law. |
| Neither characteristic loss nor observable scores improve | The selected event or transition layer does not add held-out information beyond the benchmark for this target. | Retain the reduced-form benchmark and report the generator as uninformative for this design. |
| *Score-specific evidence* | | |
| Mean scores improve, but quantile or interval scores do not | The generator captures first-moment adjustment but not the conditional distribution. | Avoid distributional claims; present the result as a mean-forecast projection. |
| Tail or rare-event scores improve while average scores do not | The adjustment law is most informative in states that are rare or nonlinear. | Emphasize tail performance and report calibration, ranking, and uncertainty intervals. |
| Direct quantile regression dominates the generator distribution | The observable distribution is sufficiently captured by direct observable histories at the chosen horizon. | Use quantile regression as the benchmark conclusion; the generator must be redesigned or applied to a different target. |

The benchmark protocol is fixed at the level of forecast targets, not at the level of model labels. Table 18 states the comparison contract. For each target, the generator supplies a forecast by taking the appropriate functional of one predictive distribution. The benchmark estimates the same observable target directly with the information set and selection rule fixed before final testing. The final score then asks whether the generator-implied distribution adds predictive information for that target, rather than whether the generator is more detailed as a representation.

Let $\mathcal{T}$ be the test block, $n_{\mathcal{T}}$ its size, and $y_i = Y_{t_i+h_i} - Y_{t_i}$ the realized observable increment for observation $i$. Let $F_{G,i}$ be the generator-implied forecast distribution for that increment, and let $F_{B,i}$ be a benchmark forecast distribution from a direct observable model, such as a vector autoregression, Gaussian transition model, Hawkes projection, or quantile regression. For a negatively oriented

Table 18: Benchmark protocol for observable distribution forecasts.

The table matches each observable forecast target to the generator-implied forecast, the reduced-form benchmark class, and the score used on the final test block. Panel headers separate low-order projections, distributional forecasts, and tail or rare-state probabilities.

| Forecast target | Generator-implied forecast | Direct benchmark | Final-test score |
|---|---|---|---|
| *Low-order observable projections* | | | |
| Conditional mean | Mean of the observable distribution induced by the selected adjustment law and measurement rule. | Vector autoregression, error-correction model, or linear projection on the same lagged observables. | Mean squared or root mean squared forecast error. |
| Conditional variance | Second central moment of the same generator-implied observable distribution. | Gaussian transition model, volatility equation, or second-moment projection. | Squared forecast error for the realized second moment or a proper density score when available. |
| *Distributional forecasts* | | | |
| Full predictive distribution | Forward transition distribution combined with the observable measurement law. | Gaussian transition density or another direct predictive distribution for the observable increment. | Observable characteristic-function loss, continuous ranked probability score approximation, or another proper distribution score. |
| Conditional quantiles | Quantiles extracted from one generator-implied predictive distribution. | Quantile regressions estimated separately or jointly for the prespecified quantile grid. | Pinball loss and quantile-crossing diagnostic. |
| Prediction interval | Central interval obtained from the lower and upper generator quantiles. | Direct interval forecast from quantile regression, Gaussian transition density, or conformal calibration fixed before testing. | Interval score, coverage, and average width. |
| *Tail and event probabilities* | | | |
| Tail probability | Probability assigned by the generator distribution to a prespecified adverse observable threshold. | Binary probability model, tail regression, or direct empirical-frequency forecast using the benchmark information set. | Brier score or squared probability loss. |
| Rare-state ranking | Generator-implied risk score for the prespecified rare state. | Hawkes event model, aggregate stress indicator, or direct classifier fixed before final testing. | Area under the receiver-operating-characteristic curve and highest-risk decile lift. |

scoring rule $S$, lower values are better, and the population comparison target is

$$\Delta_S(G, B) = \mathbb{E}[S(F_{B,i}, y_i) - S(F_{G,i}, y_i)]. \tag{30}$$

Positive values favor the generator-implied observable distribution. The sample analogue is computed on the untouched test block after the generator, the anchor, the ECF regularization, the benchmark family, the benchmark information set, and the benchmark selection rule have been fixed:

$$\hat{\Delta}_S(G, B) = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \{S(F_{B,i}, y_i) - S(F_{G,i}, y_i)\}.$$

This definition is deliberately modest: the paper does not require the generator to dominate every benchmark under every score. It asks which observable distributional features are improved by the adjustment law and which are already captured by direct reduced-form forecasts.

When an application has an explicit economic decision, the same forecast distribution enters through a decision functional. Let $\mathcal{A}_{\text{dec}}$ be the prespecified action set, let $a \in \mathcal{A}_{\text{dec}}$ be an action, and let $\ell(a, y)$ be the loss from action $a$ when the realized observable increment is $y$. The generator-implied decision is

$$a_G^*(x) \in \arg\min_{a \in \mathcal{A}_{\text{dec}}} \int \ell(a, y) F_G(\mathrm{d}y \mid x),$$

with an analogous benchmark decision $a_B^*(x)$. The held-out decision-value gap is then computed by replacing the scoring rule in (30) with the realized decision loss,

$$\Delta_\ell(G, B) = \mathbb{E}[\ell\{a_B^*(X_i), y_i\} - \ell\{a_G^*(X_i), y_i\}].$$

Positive values favor the generator for the stated decision problem. This layer is optional because many applications evaluate forecasts before specifying a policy loss. It is nevertheless important for interpretation: forecast score gains are statistical evidence about predictive distributions, while welfare or policy claims require an explicit decision loss, action set, and counterfactual or forecast regime.

The same observable forecast distribution can also be checked in the frequency domain. For a scalar observable increment, define the forecast characteristic function

$$\varphi_{m,i}^Y(v) = \int \exp(\mathrm{i}vy) F_{m,i}(\mathrm{d}y),$$

where $v$ is an observable-frequency argument. If $F_{m,i}$ is computed by simulation, this integral is replaced by the corresponding average over simulated observable draws. For an observable-frequency grid $\mathcal{U}_Y$ and standardized test functions $\omega_0, \ldots, \omega_{J_Y}$, with $\omega_0(x) \equiv 1$, the observable characteristic-function loss is

$$L_{\mathrm{OCF}}(m) = \frac{1}{|\mathcal{U}_Y|(J_Y+1)} \sum_{v \in \mathcal{U}_Y} \sum_{\ell=0}^{J_Y} \left| \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \omega_\ell(x_i)\{\exp(\mathrm{i}vy_i) - \varphi^Y_{m,i}(v)\} \right|^2 . \tag{31}$$

This is the observable analogue of (18). In the reported diagnostics, observable frequencies are standardized by the training-block standard deviation of the realized observable increment, and non-intercept conditioning functions are centered and scaled with training-block means and standard deviations before being applied to the test block. This convention makes the frequency grid comparable across observable targets and keeps final-test losses independent of test-block rescaling. The observable characteristic loss does not replace proper scoring rules. It checks whether the forecast distribution reproduces conditional characteristic restrictions for the observable itself, while pinball loss, interval scores, probability losses, and ranking scores test specific distributional features.

The scoring rules are deliberately standard, and the notation below is used consistently in the tables. The model index $m$ ranges over the generator-implied forecast and the reduced-form benchmarks; $G$ denotes the generator forecast and $B$ a benchmark when a score gap is reported. Let $F_{m,i}$ be the predictive distribution from model $m$, and

$$q_{m,i}(\tau) = \inf\{y : F_{m,i}(y) \geq \tau\}$$

the corresponding $\tau$-quantile. The reported quantile grid is ordered as $\mathcal{A} = \{\tau_1 < \cdots < \tau_{K_q}\}$. For a realized value $y_i$ and a forecasted quantile $q_{m,i}(\tau)$, the pinball loss is

$$L_\tau\{y_i, q_{m,i}(\tau)\} = \{\tau - \mathbb{1}\{y_i < q_{m,i}(\tau)\}\}\{y_i - q_{m,i}(\tau)\}. \tag{32}$$

For a quantile grid $\mathcal{A}$, the reported average pinball loss is

$$\bar{L}_m = \frac{1}{|\mathcal{A}|n_{\mathcal{T}}} \sum_{\tau \in \mathcal{A}} \sum_{i \in \mathcal{T}} L_\tau\{y_i, q_{m,i}(\tau)\}. \tag{33}$$

For a continuous predictive distribution, the CRPS can be written as twice the integral of pinball losses over all quantile levels. The reported approximation replaces that integral by the equal-weight grid average $2\bar{L}_m$. It is therefore a fixed-grid approximation used for comparable reporting, not a new scoring rule.

The reported quantile crossing share is

$$C_m = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \mathbb{1}\{q_{m,i}(\tau_{j+1}) < q_{m,i}(\tau_j) \text{ for some } j = 1, \ldots, K_q - 1\}.$$

It is a coherence diagnostic for direct quantile forecasts. A predictive distribution generated by simulation has ordered quantiles by construction after quantile extraction.

For a central $1-\alpha$ prediction interval $[\ell_{m,i,\alpha}, u_{m,i,\alpha}]$, where $\ell_{m,i,\alpha} = q_{m,i}(\alpha/2)$ and $u_{m,i,\alpha} = q_{m,i}(1-\alpha/2)$, the interval score is

$$S_{\alpha,i}(m) = u_{m,i,\alpha} - \ell_{m,i,\alpha} + \frac{2}{\alpha}\{\ell_{m,i,\alpha} - y_i\}\mathbb{1}\{y_i < \ell_{m,i,\alpha}\} + \frac{2}{\alpha}\{y_i - u_{m,i,\alpha}\}\mathbb{1}\{y_i > u_{m,i,\alpha}\}. \tag{34}$$

The reported interval coverage and width are

$$\frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \mathbb{1}\{\ell_{m,i,\alpha} \le y_i \le u_{m,i,\alpha}\}, \qquad \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} (u_{m,i,\alpha} - \ell_{m,i,\alpha}).$$

For a lower-tail threshold $c$, chosen on the training block, the reported tail probability loss is the squared probability loss

$$L_{\text{tail}}(m; c) = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \left[\mathbb{1}\{y_i \le c\} - F_{m,i}(c)\right]^2. \tag{35}$$

For binary rare-event targets, let $R_i \in \{0, 1\}$ be the realized rare-state indicator and $\pi_{m,i}$ the predicted probability. The Brier score is

$$\frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} (R_i - \pi_{m,i})^2.$$

Let $N_1 = \sum_{i \in \mathcal{T}} R_i$ and $N_0 = n_{\mathcal{T}} - N_1$. When both groups are nonempty, the area under the

receiver-operating-characteristic (ROC) curve is

$$\mathrm{AUC}_m = \frac{1}{N_1 N_0} \sum_{i \in \mathcal{T}: R_i = 1} \sum_{j \in \mathcal{T}: R_j = 0} \left[ \mathbb{1}\{\pi_{m,i} > \pi_{m,j}\} + \frac{1}{2} \mathbb{1}\{\pi_{m,i} = \pi_{m,j}\} \right]. \tag{36}$$

Thus the area-under-curve score is the probability that a randomly selected rare-state observation receives a higher score than a randomly selected non-rare observation, with ties counted as one half. Let $\mathcal{D}_m$ contain the $\lceil 0.1 n_{\mathcal{T}} \rceil$ held-out observations with the largest $\pi_{m,i}$. The decile lift is

$$\mathrm{Lift}_m = \frac{|\mathcal{D}_m|^{-1} \sum_{i \in \mathcal{D}_m} R_i}{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} R_i}.$$

The tail-concentration table also reports the full-sample rare-state rate, the highest-risk decile rare-state rate, and the highest-risk decile count:

$$\bar{R} = n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} R_i, \qquad \bar{R}_{\mathcal{D}_m} = |\mathcal{D}_m|^{-1} \sum_{i \in \mathcal{D}_m} R_i, \qquad N_{\mathcal{D}_m} = \sum_{i \in \mathcal{D}_m} R_i.$$

For scalar point forecasts $\hat{y}_{m,i}$, the root mean squared error is

$$\mathrm{RMSE}_m = \left[ n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (y_i - \hat{y}_{m,i})^2 \right]^{1/2}.$$

Benchmark gaps are reported as benchmark loss minus generator loss for negatively oriented losses and as generator score minus benchmark score for positively oriented ranking scores. Characteristic-function losses evaluate the same distributional comparison in the frequency domain. These criteria give the paper a falsifiable comparison: an estimated generator is retained as useful only where it improves or disciplines observable scores.

The score gap also clarifies the role of generator signals in ordinary regressions. A regression that adds $A_Y(x)$, $B_Y(x)$, or simulated generator quantiles to lagged observable controls is not a separate contribution. It is a projection of the same fitted distribution onto a familiar forecasting equation. If this projection improves $S$, the event or transition layer contains predictive information for the observable. If it does not, the direct observable benchmark is sufficient for that target and horizon.

The same logic applies across domains. In an order book, $Y_t$ can be the next price change, squared price change, future low-depth state, or future high-imbalance state. In a processor panel, $Y_t$ can be a payout price, delivered quantity, product-output share, inventory measure, or market-share

change. In credit, labor, or firm dynamics, $Y_t$ can be default, employment, investment, or entry outcomes. The generator enters only through the conditional adjustment signals implied by the event or transition law. The benchmark question is always the same: do those signals improve held-out observable outcomes relative to direct aggregate forecasts?

The contribution should therefore be read as a two-step claim. First, generator inversion recovers a conditional adjustment law. Second, that law is valuable only where its projections improve observable outcomes or distributional diagnostics. The simulation below follows this discipline by reporting price-change loss, volatility loss, rare-state probability loss, rare-state ranking, quantile losses, interval scores, and time-series value added alongside the characteristic-function diagnostics.

## 6 Econometric Dictionary

The generator-first view organizes familiar econometric models by the information they retain from an adjustment law. The characteristic generator in (8) is the distributional object. The entries in Table 19 are moment projections, derivatives, truncations, likelihood anchors, score projections, or direct benchmarks for that object.

This dictionary is not an equivalence theorem. A vector autoregression, volatility equation, Hawkes model, or quantile regression can be estimated without a generator. The claim is more modest and more useful: once a generator and measurement law imply an observable forecast distribution, the familiar models estimate particular functionals, lower-information moment projections, score projections, or competing forecast laws for the same observable target. The comparison is therefore empirical. A reduced-form model is sufficient when it matches the relevant observable distributional feature on held-out data; the generator adds value only when the adjustment law carries predictive information that the reduced-form information set discards.

The conditional characteristic function is the broadest diagnostic because it evaluates the generator away from $u = 0$. Moment models use only local derivatives. The first derivative gives conditional mean dynamics. The quadratic saddle approximation gives the Gaussian transition model

$$\mathrm{d}X_t = A(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \qquad \sigma(X_t)\sigma(X_t)^\top = B(X_t). \tag{37}$$

Table 19: Dictionary between generator objects and econometric representations.

The table maps the paper's generator notation to familiar econometric objects. It highlights which standard projection, volatility, tail, and persistence concepts correspond to different functionals or approximations of the generator.

| Doi–Peliti object | Econometric representation | Interpretation |
|---|---|---|
| Hamiltonian at imaginary momentum | Conditional characteristic function | Distributional law of aggregate increments |
| Quadratic saddle approximation | Autoregressive and vector error-correction projections | Linearized conditional mean dynamics |
| Second conditional cumulant | Conditional-volatility projection | State-dependent conditional variance |
| Linear self-excitation | Hawkes process | Event intensity depends on past events |
| Higher-order vertices | Non-Gaussian cumulants | Skewness, tail thickness, and asymmetric jumps |
| Instanton/minimum-action path | Rare-event probability | Most likely path to stress or depletion |
| Slow generator eigenmodes | Cointegration structures | Persistent common trends and stable spreads |

Linearizing $A(X_t)$ around an equilibrium $X^*$ and discretizing at interval $\Delta$ gives

$$X_{t+\Delta} \approx c + \Phi X_t + u_{t+\Delta}, \qquad \mathrm{Var}(u_{t+\Delta} \mid X_t) \approx \Delta B(X_t). \tag{38}$$

This is the route from a generator to autoregressive or vector-autoregressive mean dynamics and state-dependent conditional volatility. When the drift has slow stable directions and fast transitory directions, the same linearization provides a generator-based interpretation of error correction and cointegration (Johansen, 1991).

The second conditional cumulant is the volatility object. If $B(X_t)$ is persistent because the state is persistent, the discrete-time projection produces conditional heteroskedasticity in the sense of Engle (1982) and Bollerslev (1986). A Hawkes process appears when an event intensity is itself increased by past events (Hawkes, 1971). Higher vertices in (6) matter when skewness, tail thickness, or asymmetric jumps are empirically relevant. Characteristic-function losses evaluate these higher-order restrictions without estimating each cumulant separately. Rare-event probabilities use the action (7), not only the quadratic approximation.

The observable distribution $F_{Y,\theta}(\cdot \mid x)$ in (28) is the object that generates the direct forecast targets. Its conditional mean is the first-moment forecast used by autoregressive and vector-autoregressive projections. Its conditional variance is the volatility target used by conditional-volatility projections. Its inverse distribution function gives the quantiles that quantile regression estimates directly from observables. Its tail probabilities give rare-event forecasts, and its path representation gives action-based stress diagnostics. The generator is therefore not an alternative to observable modeling. It is a way to impose a common conditional law behind these observable forecasts and to test that law with scoring rules. In final testing, the reduced-form models remain direct benchmarks for the observed outcome, not automatic consequences of the generator.

The word projection is used in two distinct senses. A generator projection is a functional of the generator-implied observable distribution, such as its mean, variance, quantile, or tail probability. A direct benchmark projection is the best forecast within a restricted reduced-form class and information set under a stated scoring rule. These two objects coincide only under the corresponding correct-specification and information conditions. Otherwise their disagreement is informative: it tells the researcher whether the generator distribution, the reduced-form benchmark, or neither contains the relevant predictive information for the observable target.

Table 20: Observable forecast targets and score projections of one generator distribution.

The table shows how familiar observable forecast targets are obtained from the generator-implied predictive distribution, which reduced-form benchmark estimates the same target directly, and which score is used for held-out comparison. Panel headers separate low-order, distributional, and tail or stress targets.

| Forecast target | Generator functional | Direct benchmark | Evaluation score |
|---|---|---|---|
| *Low-order score targets* | | | |
| Conditional mean | Mean of the generator-implied observable distribution. | Autoregression, vector autoregression, or error-correction projection. | Mean-squared or root-mean-squared forecast error. |
| Conditional variance | Variance of the same generator-implied observable distribution. | Conditional-volatility equation or Gaussian transition model. | Volatility loss or distribution score. |
| *Distributional score targets* | | | |
| Conditional quantile | Inverse distribution function at a fixed probability level. | Quantile regression or another direct quantile model. | Pinball loss. |
| Prediction interval | Pair of lower and upper quantiles from one predictive distribution. | Direct interval forecast or separate quantile regressions. | Interval score, coverage, and width. |
| Full predictive distribution | Complete generator-implied law after applying the observable mapping. | Gaussian transition density, direct density forecast, or quantile-grid forecast. | Continuous-ranked-probability-score approximation, or log score when a density is available. |
| *Tail and stress score targets* | | | |
| Tail probability | Probability assigned to crossing a prespecified observable threshold. | Binary rare-event model or direct tail regression. | Probability loss, Brier score, and calibration. |
| Rare-state ranking | Ordering of held-out observations by generator-implied stress probability. | Classification or ranking model for the same rare state. | Area under the ranking curve and highest-risk decile lift. |

**Proposition 4** (Observable forecasts and reduced-form score projections)**.** *Fix a horizon $h$ and a state $x$. Suppose the estimated generator and measurement law imply the observable distribution $F_{Y,\theta}(\cdot \mid x)$. The standard forecast targets are functionals of this distribution:*

$$m_\theta(x) = \int y F_{Y,\theta}(\mathrm{d}y \mid x), \qquad v_\theta(x) = \int \{y - m_\theta(x)\}^2 F_{Y,\theta}(\mathrm{d}y \mid x),$$

$$q_{\theta,\tau}(x) = \inf\{y : F_{Y,\theta}(y \mid x) \geq \tau\}, \qquad p_{\theta,c}(x) = F_{Y,\theta}(c \mid x).$$

*The first two functionals are mean and variance forecasts, $q_{\theta,\tau}$ is the conditional quantile forecast, and $p_{\theta,c}$ is the lower-tail event probability for threshold $c$.*

*More generally, let $\mathcal{T}(F)$ be any reported forecast functional of a distribution $F$, and let $S_{\mathcal{T}}(a, y)$ be the negatively oriented scoring rule used to evaluate an action $a$ for realized outcome $y$. Squared loss elicits the mean, pinball loss elicits a quantile, interval scores evaluate central prediction intervals, probability losses evaluate threshold probabilities, and distribution scores evaluate the whole forecast law. The generator forecast is*

$$a_{G,\mathcal{T}}(x) = \mathcal{T}\{F_{Y,\theta}(\cdot \mid x)\}.$$

*If a reduced-form benchmark observes only $s(X_t)$ and restricts forecasts to a class $\mathcal{H}_{\mathcal{T}}$, its direct population target is the score projection*

$$a^*_{B,\mathcal{T}} \in \arg\min_{a \in \mathcal{H}_{\mathcal{T}}} \mathbb{E}[S_{\mathcal{T}}\{a(s(X_t)), Y_{t+h} - Y_t\}].$$

*The corresponding projection of the generator-implied law onto the same reduced information set is*

$$a^{\mathrm{proj}}_{G,\mathcal{T}} \in \arg\min_{a \in \mathcal{H}_{\mathcal{T}}} \mathbb{E}[S_{\mathcal{T}}\{a(s(X_t)), \tilde{Y}_{\theta,t+h}\}].$$

*Here $\tilde{Y}_{\theta,t+h}$ is a draw from $F_{Y,\theta}(\cdot \mid X_t)$ conditional on $X_t$. Under correct specification of the generator and measurement law, and for the same reduced information set $s(X_t)$, $a^{\mathrm{proj}}_{G,\mathcal{T}}$ and $a^*_{B,\mathcal{T}}$ coincide at the population level. Under misspecification or information loss, their disagreement is exactly what final-test score gaps evaluate.*

*Thus familiar reduced-form methods are score projections for particular forecast targets:*

$$a^*_{\mathrm{mean}} \in \arg\min_{a \in \mathcal{H}_{\mathrm{mean}}} \mathbb{E}[\{Y_{t+h} - Y_t - a(s(X_t))\}^2], \qquad a^*_\tau \in \arg\min_{a \in \mathcal{H}_\tau} \mathbb{E}[L_\tau\{Y_{t+h} - Y_t, a(s(X_t))\}].$$

*The first display covers autoregressions, vector autoregressions, and error-correction mean projections under squared loss. The second covers quantile regression under pinball loss. Analogous score projections cover Gaussian transition densities, conditional-volatility equations, and threshold-probability models. The generator enters these comparisons only through $F_{Y,\theta}$ and its score-relevant functionals.*

*Proof.* The mean, variance, quantile, and tail-probability formulas are the corresponding distributional functionals of $F_{Y,\theta}(\cdot \mid x)$. A scoring rule defines the population action that is optimal for the targeted feature. Restricting the information set to $s(X_t)$ and the forecast class to $\mathcal{H}_{\mathcal{T}}$ gives the reduced-form score projection. Replacing the realized outcome by a draw from the generator-implied forecast law gives the analogous projection of the generator distribution. If the generator-implied law is the true conditional observable law, the two expectations are taken over the same conditional distribution and have the same minimizer. If the generator is misspecified or the reduced information set discards relevant information, the two minimizers can differ. The held-out score gap compares the resulting forecasts on realized observables. □

# 7 Generated Adjustment Curves: A Demand–Supply Interpretation

This section gives one economic interpretation of generator inversion. It is not the paper's empirical claim. The link to demand-curve inversion is conceptual. A static demand curve is a market-specific object that holds latent market conditions fixed. A generated adjustment curve is a state-specific dynamic object: it describes how a demand schedule, supply schedule, inventory position, credit state, labor-flow state, or production allocation changes over time. The generator is useful only if recovering that dynamic object changes the observable distribution that the researcher wants to forecast or test.

The same logic applies once a curve or adjustment schedule is viewed as a state with a law of motion. In a processor or commodity market, the state can be a lower-frequency adjustment schedule linking product values, inventories, contracts, and benchmark prices to quantities or payout prices. In an order-driven market, it can be an aggregate of outstanding orders across price cells. The order-driven version is a transparent discrete-event implementation because arrivals and removals are observed directly, but it is only one implementation of the more general dynamic inversion argument. Let

$D_j(t)$ and $S_j(t)$ be demand and supply orders at price cell $j$. A minimal local system has four event types:

$$D_j \to D_j + 1, \qquad \text{demand arrival at rate } \lambda_j^D, \tag{39}$$

$$D_j \to D_j - 1, \qquad \text{demand cancellation at rate } \mu_j^D D_j, \tag{40}$$

$$S_j \to S_j + 1, \qquad \text{supply arrival at rate } \lambda_j^S, \tag{41}$$

$$S_j \to S_j - 1, \qquad \text{supply cancellation at rate } \mu_j^S S_j. \tag{42}$$

Adding price bins turns the local order count into demand and supply schedules. This is the sense in which arrivals and cancellations are at the core of the aggregate curves used in elementary economics.

The static curves are obtained by cumulating order stocks across price cells. If $p_j$ is a price grid, then

$$Q_t^D(p_j) = \sum_{k:p_k \geq p_j} D_k(t), \qquad Q_t^S(p_j) = \sum_{k:p_k \leq p_j} S_k(t). \tag{43}$$

The demand curve is downward sloping because fewer demand orders are willing to transact as the price rises. The supply curve is upward sloping because more supply orders are willing to transact as the price rises. The generator does not replace these curves; it supplies their law of motion. Table 21 and Figure 1 show the corresponding executable example. A sequence of arrivals and cancellations changes the order stocks, and the familiar demand and supply curves are snapshots of those stocks.

For one price cell, define excess demand $X = D - S$. Demand arrivals and supply cancellations move $X$ up by one. Demand cancellations and supply arrivals move $X$ down by one. The projected Hamiltonian is

$$H_X(D, S, \varpi) = (\lambda^D + \mu^S S)(\mathrm{e}^{\varpi} - 1) + (\mu^D D + \lambda^S)(\mathrm{e}^{-\varpi} - 1). \tag{44}$$

The conditional cumulant rates of excess demand are

$$C_k^X(D, S) = (\lambda^D + \mu^S S) + (-1)^k(\mu^D D + \lambda^S). \tag{45}$$

Thus the same event grammar produces an excess-demand drift, a state-dependent variance, higher

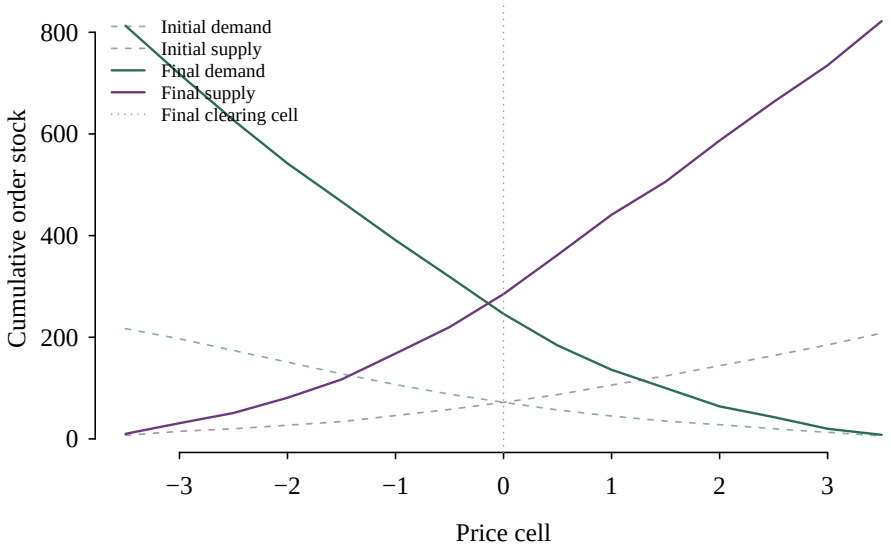
Table 21: Multi-price order-curve simulation summary.

The table reports the size and event composition of the multi-price order-curve simulation, together with initial and final clearing outcomes.

| Statistic | Value |
|---|---|
| Price cells | 15 |
| Order events | 29200 |
| Demand arrivals | 7662 |
| Demand cancellations | 7066 |
| Supply arrivals | 7543 |
| Supply cancellations | 6929 |
| Initial clearing price | 0 |
| Final clearing price | 0 |
| Clearing-price change | 0 |
| Initial clearing depth | 72 |
| Final clearing depth | 246 |

Figure 1: Demand and supply curves are cumulative snapshots of event-generated order stocks.

The figure shows cumulative demand and supply order stocks across price cells before and after the simulated sequence of arrivals and cancellations. Dashed curves are the initial stocks, solid curves are the final stocks, and the vertical guide marks the final clearing cell.



cumulants, and rare excess-demand actions.

This extension also clarifies the relation between the paper's minimal example and standard demand-supply diagrams. A static curve summarizes an order stock or adjustment schedule at a point in time. A generator describes how that object changes through observed events or lower-frequency transitions. The econometric object is not only the curve but the transition law of the curve and the observable distribution it implies.

# 8 Simulation Diagnostic for the Estimator and Score Chain

The simulation diagnostic is a verification device for the framework. It is not a domain claim. The design must be small enough to audit and rich enough to exercise the whole chain: event likelihood, regularized ECF estimation, generator-to-observable mapping, distributional forecasting, reduced-form benchmarks, and final-test score gaps. A demand–supply event grammar is useful for this purpose because the simulated events, the conditional generator, the observable projection, and the benchmark information sets are all known. Let $x_t = (D_t, S_t, p_t, \iota_t)$ collect demand orders, supply orders, price, and order imbalance $\iota_t$. At one price cell, the event set is

$$r \in \{D^+,\, D^-,\, S^+,\, S^-\},$$

where $D^+$ and $S^+$ denote demand and supply arrivals and $D^-$ and $S^-$ denote cancellations or removals. We write the event intensities as

$$\lambda_r(x_t) = e_r(x_t) \exp\left\{\theta_r^\top \phi_K(x_t)\right\}. \tag{46}$$

The exposure $e_r(x_t)$ equals one for arrivals and equals the outstanding order count for cancellations. The vector $\phi_K(x_t)$ contains basis functions. The simulation below compares linear, quadratic, spline, and Hawkes-style bases. In an empirical application, the basis should reflect the economic question: price sensitivity, inventory exposure, order imbalance, queue position, and cross-price substitution are different restrictions on the generator.

For the excess-demand stock $E_t = D_t - S_t$, the generator-implied first two cumulant rates are used

repeatedly below:

$$A_E(x_t) = \lambda_{D^+}(x_t) + \lambda_{S^-}(x_t) - \lambda_{D^-}(x_t) - \lambda_{S^+}(x_t), \tag{47}$$

$$B_E(x_t) = \lambda_{D^+}(x_t) + \lambda_{S^-}(x_t) + \lambda_{D^-}(x_t) + \lambda_{S^+}(x_t). \tag{48}$$

Here $A_E(x_t) = C_1^E(x_t)$ is the excess-demand drift rate and $B_E(x_t) = C_2^E(x_t)$ is the conditional variance rate. These two objects define the drift-error and variance-error diagnostics, the price-response slope, and the lower-order projection benchmarks.

The simulation tables use one calculation convention. The held-out block is $\mathcal{T}$, with $n_{\mathcal{T}}$ observations; hats denote fitted quantities, and subscript 0 denotes the simulated truth when it is available. Whenever a table reports an error for a known scalar generator object $G$, the number is the held-out root mean squared error

$$\left[ n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\widehat{G}_m(x_i) - G_0(x_i)\}^2 \right]^{1/2}.$$

For an observable forecast, the same calculation replaces $G_0(x_i)$ by the realized observable increment and $\widehat{G}_m(x_i)$ by the model forecast. Score gaps use the sign convention from Section 5: benchmark loss minus generator loss for negatively oriented scores, and generator score minus benchmark score for ranking measures. This convention is repeated in the tables so that a positive gap always has the same interpretation.

Each simulation result table can be traced back to the notation and calculation maps in Appendix C. The caption gives the economic object, the float description states the sample block, scale, and preferred direction, and the corresponding trace-map entry gives the formula used to compute the reported number. Thus the tables should be read as estimates of named adjustment-law, observable-distribution, or uncertainty objects, not as software outputs.

Given event times $t_i$, waiting times $\Delta_i$, pre-event states $x_i$, and event indicators $\mathbb{1}\{r_i = r\}$, the point-process log likelihood is

$$\ell_r(\theta_r) = \sum_i \mathbb{1}\{r_i = r\} \log \lambda_r(x_i) - \sum_i \lambda_r(x_i)\Delta_i. \tag{49}$$

With the log-link in (46), this can be estimated by Poisson regression with offset $\log(e_r(x_i)\Delta_i)$. This

is the first econometric object: the event likelihood uses the information that aggregate time-series projections discard.

The choice of basis is evaluated rather than assumed. This point is not cosmetic. Different representations make different approximation errors visible: generating functions, Poisson representations, spectral methods, rare-event asymptotics, and path-integral representations expose different aspects of the same master equation (Weber and Frey, 2017), while creation-operator and coherent-state representations make the stochastic-mechanics structure explicit (Baez and Biamonte, 2018). In the estimator, this means that a basis should be tied to a target. Table 22 reports a single held-out split, while Table 23 reports rolling-block validation across time. The likelihood criterion favors the Hawkes-style basis in this simulation, but the state-based linear basis has lower root-mean-square intensity error. This distinction matters. A self-exciting basis can fit event clustering without recovering the state-dependent demand and supply intensities that define the structural generator.

The two basis diagnostics are calculated as follows. For basis $b$, let $\hat{\lambda}_{r,b}(x_i)$ be the fitted intensity for event $r$ at held-out state $x_i$. The reported held-out log likelihood per interval is $n_{\mathcal{T}}^{-1} \sum_r \ell_{r,\mathcal{T}}(\hat{\theta}_{r,b})$, using (49) on the test block. Because this is a simulation, the true intensities $\lambda_{r,0}(x_i)$ are known; the reported intensity error is

$$\left[ \frac{1}{n_{\mathcal{T}} |\mathcal{R}|} \sum_{i \in \mathcal{T}} \sum_{r \in \mathcal{R}} \{ \hat{\lambda}_{r,b}(x_i) - \lambda_{r,0}(x_i) \}^2 \right]^{1/2}.$$

This intensity-error column is therefore a simulation diagnostic and would be replaced by prespecified stability, sign, external-validation, or forecast diagnostics in a real application.

Table 22: Out-of-sample fit of alternative bases for demand–supply event intensities.
The table compares candidate intensity bases on held-out event likelihood and intensity error, showing how predictive fit and structural recovery can point to different specifications.

| Basis | Test log likelihood per interval | Intensity error |
|---|---|---|
| Hawkes-style | -2.9005 | 0.9653 |
| Linear state | -2.9261 | 0.2347 |
| Spline state | -2.9313 | 0.7856 |
| Quadratic state | -2.9319 | 0.5811 |

The selection rule therefore depends on the estimand. Table 24 records the rule used in the simulation. Event prediction uses held-out likelihood. Structural recovery uses intensity error because the simulated truth is known. Aggregate generator projections use the best state-based basis, because

Table 23: Rolling-block validation of basis choice.

The table reports rolling-block validation means and dispersion for event likelihood and intensity error, indicating whether a basis choice is stable over time.

| Basis | Mean validation log likelihood | Standard deviation | Mean intensity error |
|---|---|---|---|
| Hawkes-style | -2.9347 | 0.0365 | 1.3537 |
| Linear state | -2.9663 | 0.0431 | 0.5328 |
| Quadratic state | -2.9716 | 0.0467 | 0.7113 |
| Spline state | -2.9886 | 0.0483 | 1.3644 |

the projection target is the demand–supply response to price, imbalance, and depth rather than pure event clustering. Interpretability uses the linear basis. With market data, the structural-recovery row is replaced by a prespecified diagnostic such as stability across subsamples, sign restrictions, or external validation.

Table 24: Basis choice depends on the target estimand.

The table states which basis is selected for each estimand and why the likelihood, recovery, projection, and interpretability targets need not select the same model.

| Target | Selection criterion | Basis |
|---|---|---|
| Event prediction | Highest held-out event log likelihood | Hawkes-style |
| Rolling likelihood stability | Highest mean rolling-block log likelihood | Hawkes-style |
| Structural intensity recovery | Lowest held-out true-intensity error | Linear state |
| State-based aggregate projection | Lowest held-out true-intensity error among state bases | Linear state |
| Economic interpretability | Transparent coefficients and standard errors | Linear state |

The point-process likelihood also gives a direct specification check. If $\hat{\lambda}_r$ is a good compensator for event type $r$, then the held-out martingale residual sum for event type $r$ is

$$\hat{M}_r = \sum_{i \in \mathcal{T}} \{\mathbb{1}\{r_i = r\} - \hat{\lambda}_r(x_i)\Delta_i\}$$

and should be small relative to the square root of the integrated fitted intensity

$$\hat{V}_r = \sum_{i \in \mathcal{T}} \hat{\lambda}_r(x_i)\Delta_i.$$

The standardized residual is $\hat{Z}_r = \hat{M}_r / \sqrt{\hat{V}_r}$. For the autocorrelation diagnostic, define the interval residual $\hat{m}_{r,i} = \mathbb{1}\{r_i = r\} - \hat{\lambda}_r(x_i)\Delta_i$ and its held-out mean $\bar{m}_r$. The lag-one residual autocorrelation is

$$\hat{a}_r = \frac{\sum_{i=2}^{n_\mathcal{T}} (\hat{m}_{r,i} - \bar{m}_r)(\hat{m}_{r,i-1} - \bar{m}_r)}{\sum_{i=1}^{n_\mathcal{T}} (\hat{m}_{r,i} - \bar{m}_r)^2}.$$

Table 25 reports the largest absolute $\hat{Z}_r$ across event types, the root mean square of these standardized residuals, and the largest lag-one residual autocorrelation computed within event type. This diagnostic does not prove the model is correct, but it is a useful first check before interpreting generator projections.

Table 25: Held-out martingale residual diagnostics for event intensities.

The table reports residual diagnostics for the selected event-prediction and state-projection generators, using standardized residual size, residual root mean square, and lag-one residual autocorrelation.

| Target | Basis | Largest standardized residual | Residual root mean square | Largest lag-one autocorrelation |
|---|---|---|---|---|
| Event prediction | Hawkes-style | 1.6960 | 1.4526 | 0.0271 |
| State projection | Linear state | 2.2402 | 1.3898 | 0.0276 |

A separate coarsening diagnostic asks which state information is doing the work. Table 26 compares the full order-state generator with two reduced event generators that observe less of the book. The aggregate-stock specification observes price, excess demand, and total depth, but does not use side-specific cancellation exposure. The excess-demand specification observes only price and the excess-demand stock $E_t = D_t - S_t$. The log likelihood and intensity error use the definitions above. Drift error is the held-out root mean squared difference between the fitted $A_E(x_i)$ and the true $A_{E,0}(x_i)$; variance error is the analogous root mean squared difference between $B_E(x_i)$ and $B_{E,0}(x_i)$. In this simulation, the full order state improves held-out event likelihood, intensity recovery, and generator-implied drift recovery. The variance comparison is more mixed, because total event activity can be easier to approximate than the direction and side of each event. This is the empirical content of the characteristic-generator claim: the event layer is useful only where the extra state information improves measured economic objects, not simply because the notation is richer.

Table 26: Information loss from coarsening the event state.

The table compares the full order-state generator with reduced information sets, showing how coarsening affects event likelihood, intensity recovery, drift recovery, and variance recovery.

| Information set | Log likelihood | Intensity error | Drift error | Variance error |
|---|---|---|---|---|
| Full order state | -2.9261 | 0.2347 | 0.5185 | 0.6112 |
| Aggregate stock | -2.9265 | 0.2581 | 0.5465 | 0.6113 |
| Excess demand | -2.9273 | 0.2841 | 0.5501 | 0.5451 |

The linear basis is useful for interpretation. Table 27 reports selected coefficients and standard errors. In the simulation, demand arrivals decrease with price and supply arrivals increase with price. Supply cancellations decrease with price. These signs are not imposed by the estimator; they

are recovered from event histories. In real data, this table would be a first diagnostic for whether the generator carries economically interpretable information.

Table 27: Selected linear-basis coefficients for event intensities.
The table reports price and imbalance effects for each event type under the linear state basis, with standard errors for the corresponding coefficient estimates.

| Event | Price effect | Std. error | Imbalance effect | Std. error |
|---|---|---|---|---|
| Demand arrivals | -1.0445 | 0.3101 | 1.0439 | 0.7628 |
| Demand cancellations | 0.4153 | 0.3175 | -0.0599 | 0.7662 |
| Supply arrivals | 0.8006 | 0.3091 | 0.1665 | 0.7396 |
| Supply cancellations | -0.6518 | 0.3061 | -0.4555 | 0.7611 |

The event coefficients can be aggregated into an economics-facing price-response object. Holding order stocks fixed, define

$$\partial_p A_E(x_t) = \frac{\partial A_E(x_t)}{\partial p_t},$$

the local slope of excess-demand drift with respect to price. A negative value means that a higher price lowers future excess-demand pressure through the arrival and cancellation generator. Table 28 compares the fitted slope with the simulation truth on held-out states. The slope error is $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (\partial_p \hat{A}_E(x_i) - \partial_p A_{E,0}(x_i))^2\}^{1/2}$, the slope bias is $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \{\partial_p \hat{A}_E(x_i) - \partial_p A_{E,0}(x_i)\}$, and sign agreement is the share of held-out states for which the fitted and true slopes have the same sign. The fitted generator recovers the negative sign in all held-out states, but it overstates the magnitude of the response. This diagnostic therefore connects the event estimator to the basic demand-supply restriction while making clear that economic sign recovery is weaker than full recovery of the price-response function.

Table 28: Generator-implied price response of excess-demand drift.
The table compares the true and estimated price-response slopes of excess-demand drift on held-out states, including sign recovery and slope-error diagnostics.

| Price-response diagnostic | Value |
|---|---|
| Selected state basis | Linear state |
| Median true slope | -10.4122 |
| Median estimated slope | -14.0707 |
| Slope error | 3.7962 |
| Slope bias | -3.7709 |
| True negative share | 100% |
| Estimated negative share | 100% |
| Sign agreement | 100% |
| Held-out states | 2006 |

The second object is the projection to observable aggregate dynamics. Price adjustment and volatility can be modeled as projections of $A_E(x_t)$ and $B_E(x_t)$, rather than as unrestricted reduced-form equations. The same logic applies to delivered quantities, inventory changes, product shares, or stress labels in lower-frequency markets. This creates direct tests against aggregate mean, volatility, and event-clustering benchmarks. In the simulation, the first conditional-moment check asks whether $A_E(x_i)\Delta_i$ predicts the next excess-demand increment $\Delta E_i$ better than an aggregate autoregression. The second check asks whether $B_E(x_i)\Delta_i + \{A_E(x_i)\Delta_i\}^2$ fits the next squared increment $(\Delta E_i)^2$ better than an aggregate second-moment regression. The errors in Table 29 are held-out root mean squared errors. For example, the excess-demand drift error is $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (\Delta E_i - \hat{m}_i)^2\}^{1/2}$, where $\hat{m}_i$ is either the generator-implied or benchmark-predicted mean increment. The second-moment error replaces $\Delta E_i$ by $(\Delta E_i)^2$ and $\hat{m}_i$ by the predicted second moment; the price-change and squared-price-change errors use the same rule with the observable price increment.

Table 29: Projection benchmarks in the demand–supply simulation.

The table compares generator-based projections with aggregate reduced-form benchmarks for conditional moments, price adjustment, and volatility.

| Model | Test loss |
|---|---|
| *Conditional moments: Excess-demand drift error* | |
| Generator signals | 2.2148 |
| Time-series benchmark | 2.2024 |
| *Conditional moments: Second-moment error* | |
| Generator signals | 7.6876 |
| Time-series benchmark | 7.7371 |
| *Price adjustment: Price-change error* | |
| Generator signals | 0.0152 |
| Vector-autoregression benchmark | 0.0223 |
| *Volatility: Squared price-change error* | |
| Generator signals | 0.0010 |
| Conditional-volatility benchmark | 0.0010 |

The conditional-moment results are intentionally reported even when they are not favorable to the generator. In the current draw, the aggregate regression slightly improves the first-moment fit, while the generator slightly improves the second-moment fit. This is the diagnostic role of the exercise: it shows where the estimated event generator carries information and where an aggregate projection is already sufficient.

The characteristic-generator diagnostic evaluates the same models as distributional objects rather than only through means and variances. The likelihood-fitted generator supplies the initialization

and benchmark. The regularized ECF generator then refines the same state-basis coefficients by minimizing the stacked characteristic moments in (20), subject to validation against the likelihood anchor. For excess-demand increments, a jump generator implies

$$\phi_{\hat{\theta}}(u \mid x_t, \Delta) = \exp\left\{\Delta\left[\hat{\lambda}_+(x_t)(e^{\mathrm{i}u} - 1) + \hat{\lambda}_-(x_t)(e^{-\mathrm{i}u} - 1)\right]\right\},$$

where $\hat{\lambda}_+ = \hat{\lambda}_{D^+} + \hat{\lambda}_{S^-}$ and $\hat{\lambda}_- = \hat{\lambda}_{D^-} + \hat{\lambda}_{S^+}$. Table 30 reports the held-out conditional-characteristic-function loss in (18). In this simulation, the low band contains 0.25, 0.50, and 0.75; the medium band contains 1.00, 1.50, and 2.00; and the high band contains 2.50, 2.85, and 3.10. The conditional loss uses an intercept and four state instruments: excess demand, pre-event price, pre-event imbalance, and log pre-event depth. Non-intercept instruments are centered and scaled with training-block means and standard deviations, and the same transformation is then applied on the test block. The unconditional loss repeats the same frequency calculation with only the intercept instrument. The oracle jump generator is included only because the data are simulated. It reports the finite-sample loss of the true conditional characteristic function and therefore gives a benchmark for sampling noise. The likelihood generator improves the low-frequency conditional loss relative to the aggregate Gaussian and Hawkes-style alternatives. The regularized ECF step selects the likelihood anchor in this calibration, avoiding the deterioration produced by an unconstrained ECF refit. The high-frequency losses are close across models, which cautions against treating every frequency as equally informative in a finite simulated sample.

The same logic can be used with an ordinary real time series. A forecaster may first estimate a price-change equation, a volatility equation, or a stress-probability equation using only lagged aggregate prices, depth, and imbalance. If event data are available, the estimated generator adds order-flow signals: excess-demand pressure, total event activity, and depth pressure. Table 31 reports the held-out value of adding those signals to the aggregate time-series benchmark. For root mean squared errors and Brier scores, the gain is the time-series-only loss minus the loss after adding generator signals. For the ranking row, the gain is the area under the receiver-operating-characteristic curve after adding generator signals minus the time-series-only ranking score. Thus a positive value always favors adding generator signals.

The observable distribution forecast is a stronger test. For each held-out state, the fitted generator is used to simulate forward event counts over the next interval. Specifically, for simulation path $s$,

Table 30: Conditional characteristic-function diagnostics for excess-demand increments, with losses scaled by $10^4$.

The table reports conditional and unconditional characteristic-function losses by frequency band, comparing oracle, likelihood, regularized ECF, Gaussian, aggregate, and Hawkes-style generators. Entries equal $10^4$ times the loss in (18); the oracle-distance column is scaled in the same way. Conditional losses use the standardized training-block instruments described in the text, while unconditional losses use only the intercept. The regularized ECF row is validation-adaptive and coincides with the likelihood generator when the held-out characteristic loss does not support moving away from the likelihood anchor.

| Model | Conditional characteristic loss | Unconditional characteristic loss | Oracle characteristic distance |
|---|---|---|---|
| *Low (0.25–0.75)* | | | |
| Oracle jump generator | 12.758 | 5.899 | 0.000 |
| Estimated jump generator | 15.784 | 6.381 | 8.898 |
| Regularized characteristic-function jump generator | 15.784 | 6.381 | 8.898 |
| Gaussian truncation | 16.217 | 9.575 | 13.549 |
| Aggregate Gaussian projection | 21.916 | 18.529 | 36.724 |
| Hawkes-style generator | 56.518 | 4.126 | 83.748 |
| *Medium (1.00–2.00)* | | | |
| Oracle jump generator | 20.005 | 31.751 | 0.000 |
| Estimated jump generator | 21.223 | 32.389 | 0.662 |
| Regularized characteristic-function jump generator | 21.223 | 32.389 | 0.662 |
| Gaussian truncation | 21.594 | 35.076 | 3.416 |
| Aggregate Gaussian projection | 21.221 | 34.290 | 8.663 |
| Hawkes-style generator | 24.441 | 31.203 | 9.492 |
| *High (2.50–3.10)* | | | |
| Oracle jump generator | 10.796 | 12.969 | 0.000 |
| Estimated jump generator | 10.799 | 12.974 | 0.000 |
| Regularized characteristic-function jump generator | 10.799 | 12.974 | 0.000 |
| Gaussian truncation | 10.809 | 12.989 | 0.000 |
| Aggregate Gaussian projection | 10.808 | 12.988 | 0.000 |
| Hawkes-style generator | 10.772 | 12.946 | 0.002 |

Table 31: Value added by generator signals in aggregate time-series forecasts.

The table reports held-out forecast performance with and without generator signals for price changes, squared price changes, rare-stress probability, and rare-stress ranking. Entries are final-test scores. The generator-gain column is computed as benchmark loss minus augmented loss for negatively oriented scores and as augmented ranking score minus benchmark ranking score for the area-under-curve row.

| Forecast target | Loss or score | Time series only | With generator signals | Generator gain |
|---|---|---|---|---|
| Price changes | Root mean squared error | 0.0223 | 0.0060 | 0.0163 |
| Squared price changes | Root mean squared error | 0.0010 | 0.0010 | 0.0000 |
| Rare-stress probability | Brier score | 0.1519 | 0.1477 | 0.0042 |
| Rare-stress ranking | Area under receiver-operating-characteristic curve | 0.6616 | 0.6504 | -0.0112 |

the four event counts are drawn from independent Poisson laws with means $\hat{\lambda}_r(x_i)\Delta_i$. The simulated excess-demand increment is

$$\Delta E_i^{(s)} = N_{D^+,i}^{(s)} + N_{S^-,i}^{(s)} - N_{D^-,i}^{(s)} - N_{S^+,i}^{(s)}.$$

Let $b_{\hat{\beta}}$ denote the linear price-change measurement law estimated by least squares on the training block. Its dependent variable is the next price change. Its regressors are the simulated excess-demand increment, price $p_i$, imbalance $\iota_i$, log displayed depth $\log Q_i$, with $Q_i = D_i + S_i$, and the generator-implied excess-demand pressure and event-activity signals. If $\hat{\varepsilon}^{(s)}$ is drawn with replacement from the training residuals of that measurement law, the simulated observable draw is

$$\tilde{y}_i^{(s)} = b_{\hat{\beta}}\{\Delta E_i^{(s)}, p_i, \iota_i, \log Q_i, \widehat{A}_E(x_i), \widehat{B}_E(x_i)\} + \hat{\varepsilon}^{(s)}.$$

The empirical distribution of these draws is the generator forecast in (29). It is compared with two reduced-form quantile regressions estimated on the training block: one using aggregate state variables and one adding generator-implied signals. The reported quantiles are 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, and 0.95. The lower-tail probability loss uses the tenth percentile of training-block price changes as the threshold. The observable characteristic loss uses the standardized observable-frequency grid described in Section 5 and the training-standardized instruments price, imbalance, and log depth. Table 32 reports summary distribution scores. In this calibration the generator distribution is competitive but does not dominate the aggregate quantile-regression benchmark on average pinball loss. It does better on the observable characteristic-function loss, which checks the whole forecast transform rather than selected quantiles. Its interval coverage is close to nominal, and the 90 percent interval score is slightly lower than the aggregate quantile-regression score. This is the right empirical standard: the generator must earn its place by forecasting the distribution of an observable, not by being richer.

Table 33 reports the same comparison as score gaps against each direct distributional benchmark and adds moving-block bootstrap uncertainty. This makes the value claim explicit: a positive entry means that the generator-implied observable distribution has lower final-test loss for the stated criterion, while a negative entry means that the direct benchmark is better for that feature of the price-change distribution. The bootstrap intervals resample held-out prediction rows; they quantify evaluation uncertainty for the fixed fitted models rather than full first-stage estimation uncertainty.

Table 32: Observable distribution forecast scores for price changes.

The table compares generator-implied predictive distributions with aggregate quantile regression and quantile regression augmented by generator signals, using average pinball loss, a quantile-grid approximation to the continuous ranked probability score, observable characteristic-function loss, downside-tail probability loss, quantile-crossing share, and the number of generator simulation paths. All losses are final-test averages. The CRPS approximation equals twice the average pinball loss over the displayed quantile grid. The observable characteristic loss is multiplied by $10^4$; for direct quantile benchmarks it uses the displayed quantile grid as a finite forecast-distribution approximation. The downside-tail threshold is the tenth percentile of training-block price changes.

| Model | Average pinball loss | Continuous ranked probability score approximation | Observable characteristic-function loss $\times 10^4$ | Tail probability loss | Crossing share | Simulation paths |
|---|---|---|---|---|---|---|
| Generator distribution | 0.0050 | 0.0099 | 24.0191 | 0.0696 | 0% | 250 |
| Aggregate quantile regression | 0.0050 | 0.0099 | 161.8189 | 0.0695 | 0% | – |
| Quantile regression with generator signals | 0.0050 | 0.0101 | 396.3519 | 0.0755 | 0% | – |

Table 33: Observable distribution score gaps for price changes.

The table converts the distribution-score levels into final-test gaps against each reduced-form benchmark and reports moving-block bootstrap intervals with positive-gap shares. Each panel fixes the direct benchmark, and entries equal benchmark loss minus generator loss, so positive values favor the generator-implied observable distribution. Observable characteristic-function losses are scaled by $10^4$, matching Table 32.

| Criterion | Generator gap | 95% interval | Positive share |
|---|---|---|---|
| *Benchmark: Aggregate quantile regression* | | | |
| Average pinball loss | -1.90e-06 | [-6.02e-05, 4.48e-05] | 41% |
| Continuous ranked probability score approximation | -3.80e-06 | [-1.20e-04, 8.96e-05] | 41% |
| Downside-tail probability loss | -1.01e-04 | [-0.0015, 0.0016] | 48% |
| Observable characteristic-function loss $\times 10^4$ | 137.7998 | [75.6816, 203.3549] | 100% |
| *Benchmark: Quantile regression with generator signals* | | | |
| Average pinball loss | 7.43e-05 | [-2.06e-05, 1.54e-04] | 90% |
| Continuous ranked probability score approximation | 1.49e-04 | [-4.12e-05, 3.08e-04] | 90% |
| Downside-tail probability loss | 0.0059 | [0.0024, 0.0086] | 100% |
| Observable characteristic-function loss $\times 10^4$ | 372.3328 | [250.0791, 564.7534] | 100% |

Table 34: Quantile forecast accuracy for observable price changes.

The table reports final-test pinball loss at each forecasted quantile in the prespecified grid $\{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$, allowing the generator-implied distribution to be compared with direct quantile-regression forecasts across the center and tails of the price-change distribution.

| Quantile | Generator distribution | Aggregate quantile regression | Quantile regression with generator signals |
|---|---|---|---|
| 0.05 | 0.0025 | 0.0026 | 0.0025 |
| 0.10 | 0.0040 | 0.0040 | 0.0042 |
| 0.25 | 0.0069 | 0.0069 | 0.0070 |
| 0.50 | 0.0084 | 0.0084 | 0.0084 |
| 0.75 | 0.0067 | 0.0066 | 0.0068 |
| 0.90 | 0.0038 | 0.0038 | 0.0039 |
| 0.95 | 0.0024 | 0.0024 | 0.0024 |

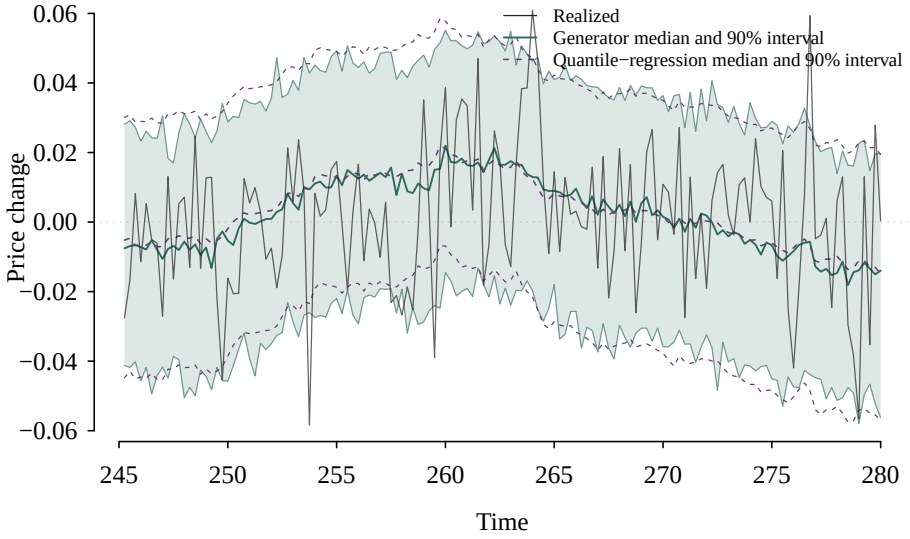
Table 35: Prediction interval calibration for observable price changes.

The table reports final-test coverage, average width, and interval scores for 80 percent and 90 percent prediction intervals. The 80 percent interval uses the 0.10 and 0.90 quantiles; the 90 percent interval uses the 0.05 and 0.95 quantiles. Panel headers separate the interval levels so the model names do not repeat the same interval label down the table.

| Model | Coverage | Average width | Interval score | Observations |
|---|---|---|---|---|
| *80 percent interval* | | | | |
| Generator distribution | 0.8091 | 0.0534 | 0.0789 | 419 |
| Aggregate quantile regression | 0.8091 | 0.0531 | 0.0785 | 419 |
| Quantile regression with generator signals | 0.8640 | 0.0632 | 0.0813 | 419 |
| *90 percent interval* | | | | |
| Generator distribution | 0.8974 | 0.0707 | 0.0981 | 419 |
| Aggregate quantile regression | 0.9021 | 0.0722 | 0.0988 | 419 |
| Quantile regression with generator signals | 0.9260 | 0.0792 | 0.0993 | 419 |

Figure 2: Generator and quantile-regression prediction intervals for price changes.

The plot shows a held-out segment of realized price changes together with the generator-implied median and 90 percent interval and the aggregate quantile-regression median and 90 percent interval. The generator interval is shown as a light band; the quantile-regression interval is shown with dashed bounds.



The characteristic generator also identifies which cumulants Gaussian aggregate projections discard by construction. The notation extends the earlier definitions $A_E = C_1^E$ and $B_E = C_2^E$. For unit jumps in excess demand, the third cumulant rate is

$$C_3^E(x_t) = \lambda_{D^+}(x_t) + \lambda_{S^-}(x_t) - \lambda_{D^-}(x_t) - \lambda_{S^+}(x_t),$$

and the fourth cumulant rate is

$$C_4^E(x_t) = \lambda_{D^+}(x_t) + \lambda_{S^-}(x_t) + \lambda_{D^-}(x_t) + \lambda_{S^+}(x_t).$$

A Gaussian diffusion, vector-autoregressive, or conditional-volatility projection uses the first two conditional cumulants and sets the higher cumulants to zero. Table 36 and Figure 3 evaluate this loss directly in the simulation by comparing generator-implied third and fourth cumulants with the true event-generator cumulants on held-out intervals. The increment-error entries are root mean squared errors for interval cumulants $C_k^E(x_i)\Delta_i$. The standardized-skewness entry reports

$$n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \left| \frac{C_3^E(x_i)\Delta_i}{\{C_2^E(x_i)\Delta_i\}^{3/2}} \right|,$$

and the excess-kurtosis entry reports the held-out median of

$$\frac{C_4^E(x_i)\Delta_i}{\{C_2^E(x_i)\Delta_i\}^2}.$$

Table 36: Higher-cumulant diagnostics for excess-demand increments.

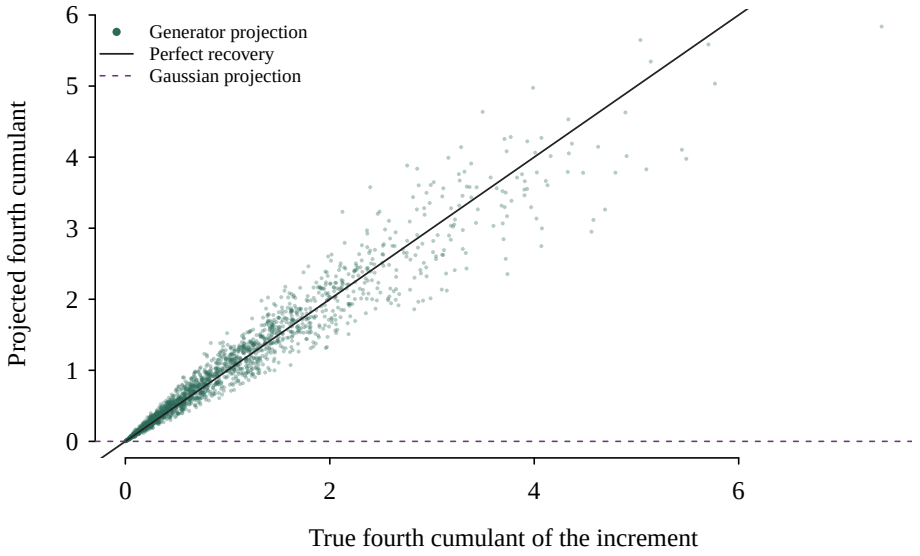
The table compares generator and Gaussian projections for third and fourth cumulants, standardized skewness, and excess kurtosis of excess-demand increments.

| Target | Model | Metric | Value |
|---|---|---|---|
| Third cumulant | Generator signals | Increment error | 0.0978 |
| Third cumulant | Gaussian projection | Increment error | 0.1144 |
| Fourth cumulant | Generator signals | Increment error | 0.2292 |
| Fourth cumulant | Gaussian projection | Increment error | 1.3686 |
| Standardized skewness | Generator signals | Mean absolute value | 0.0572 |
| Excess kurtosis | Generator signals | Median | 1.3680 |

Rare-event prediction is evaluated separately. A rare liquidity state is defined from the training sample: a held-out observation is labeled rare if the forward 20-event horizon reaches either a low-depth state below the training-sample 10th percentile of total depth or an extreme-imbalance

Figure 3: Generator projections retain fourth-cumulant information that Gaussian projections set to zero.

Each point is a held-out interval. The horizontal axis reports the true fourth cumulant of the excess-demand increment, while the vertical axis reports the generator projection. The diagonal marks perfect recovery; the horizontal reference shows the Gaussian projection, which sets higher cumulants to zero.




state above the training-sample 90th percentile of absolute imbalance. Table 37 compares a logit using generator-implied drift and variance features with a logit using aggregate lagged variables. The simulation provides weak evidence on ranking, with similar area-under-curve scores, but the generator features improve the probability loss. This is exactly the kind of result the empirical paper must test on real event data rather than assert from theory.

Table 37: Rare liquidity-state prediction in the simulation.

The table reports final-test probability and ranking scores for rare liquidity-state classification under generator and aggregate rare-event models. Lower Brier scores and higher area-under-curve scores are better. The rare-state label uses the training-sample depth and imbalance cutoffs described in the text.

| Model | Metric | Test value |
|---|---|---|
| Generator rare-event model | Area under receiver-operating-characteristic curve | 0.6628 |
| Time-series rare-event model | Area under receiver-operating-characteristic curve | 0.6616 |
| Generator rare-event model | Brier score | 0.1172 |
| Time-series rare-event model | Brier score | 0.1519 |

Area-under-curve and Brier scores summarize average performance, but rare breakdowns also require a tail check. Table 38 asks whether the highest predicted-risk decile actually contains a higher breakdown rate than the held-out sample as a whole. The reported breakdown rate is the held-out

mean of the rare-state indicator, the highest-risk decile breakdown rate is the same mean restricted to the highest predicted-risk decile, and the decile positives are the corresponding count. This diagnostic is deliberately direct: a useful stress signal should concentrate future low-depth or high-imbalance states in the observations it ranks as most fragile.

Table 38: Tail concentration of rare liquidity-state forecasts.

The table shows whether the highest predicted-risk decile contains more future rare liquidity states than the held-out sample as a whole. The breakdown rate is the held-out rare-state share, the highest-risk decile rate is the same share restricted to the 10 percent of observations with the largest predicted risk, and lift is the ratio of the two rates.

| Model | Tail diagnostic | Value |
|---|---|---|
| Generator rare-event model | Held-out observations | 420 |
| Generator rare-event model | Breakdown rate | 19.0% |
| Generator rare-event model | Highest-risk decile breakdown rate | 78.6% |
| Generator rare-event model | Highest-risk decile lift | 4.1250 |
| Generator rare-event model | Highest-risk decile positives | 33 |
| Time-series rare-event model | Held-out observations | 420 |
| Time-series rare-event model | Breakdown rate | 19.0% |
| Time-series rare-event model | Highest-risk decile breakdown rate | 14.3% |
| Time-series rare-event model | Highest-risk decile lift | 0.7500 |
| Time-series rare-event model | Highest-risk decile positives | 6 |

Rare-event actions ask a different question from classification. The logit benchmark predicts whether a stress label occurs. The action diagnostic asks how costly a path to stress is under the estimated jump generator. For displayed depth $Q_t = D_t + S_t$, define the depth-increasing rate

$$\lambda_+(x_t) = \lambda_{D^+}(x_t) + \lambda_{S^+}(x_t)$$

and the depth-decreasing rate

$$\lambda_-(x_t) = \lambda_{D^-}(x_t) + \lambda_{S^-}(x_t).$$

The local jump action for a depth velocity $v$ is

$$L(v; x_t) = \sup_{\varpi} \left[ \varpi v - \lambda_+(x_t)(e^{\varpi} - 1) - \lambda_-(x_t)(e^{-\varpi} - 1) \right].$$

For held-out state $i$, let $q_i = D_i + S_i$ be displayed depth, let $q_0$ be the training-sample low-depth cutoff, let $h_{\text{act}}$ be the fixed action horizon, and set $q_i^* = \min\{q_i, q_0\}$. The diagnostic path is the straight line from $q_i$ to $q_i^*$, with constant velocity $v_i = (q_i^* - q_i)/h_{\text{act}}$. The reported fitted jump action is

$$\hat{I}_i = h_{\text{act}} L(v_i; x_i),$$

with $\lambda_+$ and $\lambda_-$ replaced by their fitted values. The true jump action and the Gaussian action use the same path and horizon, replacing the fitted jump rates by the simulated truth or by the quadratic Gaussian local action. Table 39 and Figure 4 evaluate a fixed-horizon linear path from each held-out state to the training-sample low-depth threshold. The comparison is local and diagnostic. The root mean squared error is computed relative to the true jump action. The action ratio is the candidate action divided by the true jump action, reported as a held-out median. If $\hat{I}_i$ is a candidate action and $I_{0,i}$ is the true jump action for held-out state $i$, the rank-correlation entry is

$$\mathrm{Corr}\{\mathrm{rank}(\hat{I}_i), \mathrm{rank}(I_{0,i})\},$$

the Spearman correlation between candidate and true action values across held-out states. In this calibration, the true Gaussian truncation is close to the true jump action for the chosen path, while the estimated jump and estimated Gaussian actions mainly reflect first-stage intensity error. This is a useful negative result: rare-event action diagnostics should test whether the full jump geometry matters, not assume it matters in every state and horizon. Proposition 16 treats this fixed-path plug-in object as a smooth generator projection; optimized rare-event probabilities remain a separate empirical and theoretical problem.

Table 39: Rare-event action diagnostics for depth depletion.

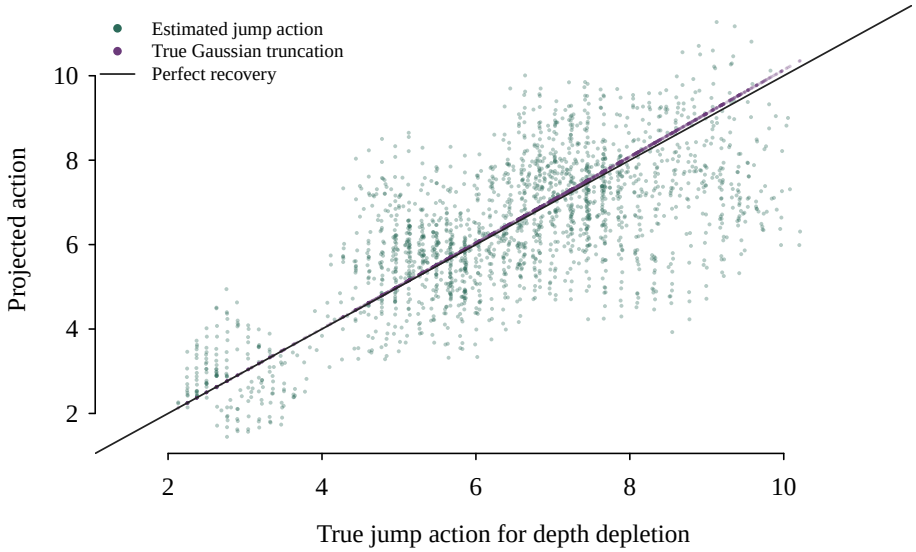
The table compares estimated and benchmark action values for fixed-horizon paths to the low-depth threshold, using error, action-ratio, and rank-correlation diagnostics.

| Model | Metric | Value |
|---|---|---|
| Threshold: 224.0000; horizon: 8.0000; held-out states: 2006. | | |
| Estimated jump action | Root mean squared error relative to true jump action | 1.3418 |
| Estimated Gaussian action | Root mean squared error relative to true jump action | 1.3348 |
| True Gaussian truncation | Root mean squared error relative to true jump action | 0.0600 |
| Estimated jump action | Median action ratio | 0.9920 |
| Estimated Gaussian action | Median action ratio | 0.9991 |
| True Gaussian truncation | Median action ratio | 1.0071 |
| Estimated jump action | Spearman rank correlation | 0.6342 |
| Estimated Gaussian action | Spearman rank correlation | 0.6415 |
| True Gaussian truncation | Spearman rank correlation | 1.0000 |

The benchmark gaps also need uncertainty. Table 40 reports a block bootstrap over held-out prediction rows. A positive gap means that the generator has lower loss than the aggregate benchmark, except for the rare-event ranking score where it means a higher value. The reported interval is the 2.5 and 97.5 percentiles of the bootstrap gap distribution, and the positive share is the fraction of bootstrap draws in which the gap is above zero. This is not a full bootstrap of the first-stage gen-

Figure 4: Estimated jump actions and Gaussian actions for paths to the low-depth threshold.

The scatter plot compares fixed-horizon depth-depletion costs across held-out states. The horizontal axis is the true jump action, while the vertical axis reports estimated jump actions and Gaussian action values; the diagonal is exact agreement with the true jump-action benchmark.




erator estimator. It is an evaluation uncertainty check for the held-out comparison. In the current simulation, the price-adjustment gain is stable, the volatility comparison favors the conditional-volatility benchmark, and the rare-event probability gain is positive but imprecise.

A stronger diagnostic resamples before the generator is estimated. Tables 41 and 42 repeat the state-basis selection and the first-stage intensity estimation inside each block-bootstrap draw, then recompute the aggregate projection gaps. The validation rule selects among the state-based bases by event likelihood before the final test comparison is evaluated. This does not prove bootstrap validity for the sieve estimator, but it moves the uncertainty calculation to the right level: basis choice, event-intensity estimation, and aggregate projection are all treated as estimated objects.

The block length is part of the inferential design. Table 43 repeats the first-stage bootstrap at several block lengths and reports whether the sign and width of the projection-gap intervals are stable. The interval width is the upper endpoint minus the lower endpoint of the bootstrap interval. This is a diagnostic rather than a theorem: a real empirical application still has to justify the block length from market persistence and sampling frequency. It is nevertheless useful because it makes the sensitivity of the uncertainty calculation observable.

Table 40: Block-bootstrap uncertainty for held-out benchmark gaps.

The table reports moving-block bootstrap intervals for held-out benchmark gaps, separating stable gains from comparisons whose sign remains uncertain. The current build uses 500 bootstrap draws and block length 8 for this held-out-row resampling. Positive generator gaps favor the generator under the stated loss or ranking score.

| Criterion | Generator gap | 95% interval | Positive share |
|---|---|---|---|
| *Conditional moments* | | | |
| Excess-demand drift error | -0.0124 | [-0.0215, 0.0037] | 7% |
| Second-moment error | 0.0496 | [-0.0273, 0.1366] | 90% |
| *Price adjustment* | | | |
| Price-change error | 0.0071 | [0.0051, 0.0084] | 100% |
| *Volatility* | | | |
| Squared price-change error | 0.0000 | [-0.0001, 0.0000] | 0% |
| *Rare liquidity state* | | | |
| Probability loss | 0.0347 | [-0.0111, 0.0706] | 92% |
| Ranking score | 0.0012 | [-0.2597, 0.2263] | 49% |

Table 41: Basis-selection stability in the first-stage bootstrap.

The table reports how often each state basis is selected when basis choice is repeated inside the first-stage block bootstrap.

| Basis | Selection share | Count | Replications |
|---|---|---|---|
| Linear state | 100% | 40 | 40 |

Table 42: First-stage block-bootstrap uncertainty for projection gaps.

The table reports benchmark-gap uncertainty when basis selection and event-intensity estimation are repeated inside each bootstrap draw. The current build uses 40 first-stage bootstrap draws and block length 80. Positive generator gaps favor the generator under the stated loss or ranking score.

| Criterion | Generator gap | 95% interval | Positive share |
|---|---|---|---|
| *Conditional moments* | | | |
| Excess-demand drift error | -0.0124 | [-0.0705, 0.0012] | 8% |
| Second-moment error | 0.0496 | [0.0139, 0.0611] | 100% |
| *Price adjustment* | | | |
| Price-change error | 0.0071 | [0.0014, 0.0143] | 100% |
| *Volatility* | | | |
| Squared price-change error | 0.0000 | [-0.0001, 0.0000] | 0% |
| *Rare liquidity state* | | | |
| Probability loss | 0.0347 | [-0.0034, 0.0694] | 90% |
| Ranking score | 0.0012 | [-0.0497, 0.2386] | 80% |

Table 43: Block-length sensitivity for first-stage bootstrap gaps.

The table compares first-stage bootstrap gaps across block lengths 40, 80, and 160, showing whether interval width and sign are sensitive to the dependence window. Panel headers state the target and criterion; positive generator gaps favor the generator.

| Block length | Generator gap | Interval width | Positive share |
|---|---|---|---|
| *Conditional moments: Excess-demand drift error* | | | |
| 40 | -0.0124 | 0.0409 | 0% |
| 80 | -0.0124 | 0.0567 | 8% |
| 160 | -0.0124 | 0.0429 | 8% |
| *Conditional moments: Second-moment error* | | | |
| 40 | 0.0496 | 0.0222 | 100% |
| 80 | 0.0496 | 0.0326 | 100% |
| 160 | 0.0496 | 0.0233 | 100% |
| *Price adjustment: Price-change error* | | | |
| 40 | 0.0071 | 0.0098 | 100% |
| 80 | 0.0071 | 0.0130 | 100% |
| 160 | 0.0071 | 0.0134 | 100% |
| *Rare liquidity state: Ranking score* | | | |
| 40 | 0.0012 | 0.2344 | 83% |
| 80 | 0.0012 | 0.2390 | 83% |
| 160 | 0.0012 | 0.1744 | 75% |
| *Rare liquidity state: Probability loss* | | | |
| 40 | 0.0347 | 0.0570 | 100% |
| 80 | 0.0347 | 0.0725 | 83% |
| 160 | 0.0347 | 0.0462 | 83% |
| *Volatility: Squared price-change error* | | | |
| 40 | 0.0000 | 0.0001 | 0% |
| 80 | 0.0000 | 0.0001 | 0% |
| 160 | 0.0000 | 0.0001 | 0% |

A single simulation draw is too weak for a general empirical claim. Table 44 therefore repeats the complete simulation, estimation, basis selection, and benchmark comparison across independent replications. Ratios divide the generator-based loss by the corresponding benchmark loss, so values below one favor the generator-based object. Differences subtract the benchmark score from the generator score, so values above zero favor the generator-based object. The generator-favored share is the share of finite replications satisfying the relevant inequality. The replication count is criterion-specific because rare-state scores are defined only when the held-out block contains both rare and non-rare states. A pilot table checks the workflow, while a submission version should use at least 100 finite observations for every reported criterion. The Monte Carlo exercise is not a substitute for market data, but it checks whether the result depends on one path. It also separates two questions that are easy to conflate. A Hawkes-style basis can be competitive for event likelihood because it captures clustering, while a state-based demand–supply basis can recover the structural intensity surface more accurately and produce better aggregate projections.

Table 44: Monte Carlo stability of generator benchmarks.

The table summarizes repeated simulation, estimation, basis selection, and benchmark evaluation across independent replications, reporting medians, interquartile ranges, and the share favoring the generator. Ratios below one favor the generator; score differences above zero favor the generator. Replication counts are criterion-specific, so rare-state scores can use fewer finite replications when a held-out block does not contain both classes.

| Criterion | Median | Interquartile range | Generator-favored share | Replications |
|---|---|---|---|---|
| *Generator recovery* | | | | |
| Intensity error ratio | 0.1921 | [0.1811, 0.2179] | 100% | 6 |
| *Event likelihood* | | | | |
| Log likelihood difference | -0.0277 | [-0.0290, -0.0250] | 0% | 6 |
| *Conditional moments* | | | | |
| Drift error ratio | 1.0136 | [1.0115, 1.0177] | 0% | 6 |
| Second-moment error ratio | 0.9980 | [0.9948, 1.0007] | 67% | 6 |
| *Price adjustment* | | | | |
| Price-change error ratio | 0.8858 | [0.8499, 0.9082] | 100% | 6 |
| *Volatility* | | | | |
| Volatility error ratio | 1.0123 | [1.0098, 1.0151] | 0% | 6 |
| *Rare liquidity* | | | | |
| Probability-loss ratio | 0.8943 | [0.8433, 0.9823] | 67% | 6 |
| Ranking-score difference | 0.0432 | [-0.0767, 0.1738] | 67% | 6 |

Current build reports pilot Monte Carlo evidence; final submission evidence should use at least 100 finite criterion-specific Monte Carlo observations for every reported criterion.

The simulation evidence shows the identification logic and the benchmark structure. A market-data application would replace or complement the simulated event history with observed order messages, prespecified basis families, out-of-sample likelihood and forecast evaluation, and rare-event validation.

Figure 5: Basis choice creates a prediction–recovery tradeoff.

The plot places each candidate basis by held-out event likelihood and true-intensity recovery error. It shows that the Hawkes-style basis predicts events best in this simulation, while the linear state basis better recovers the structural intensity surface.

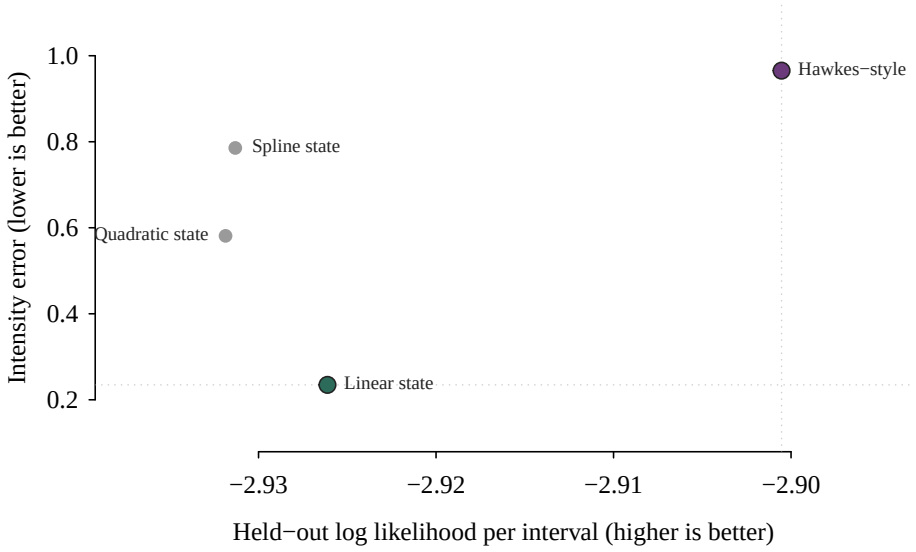



Figure 6: Rolling-block validation checks whether basis performance is stable over time.

The figure summarizes rolling-block validation across time blocks. Points show mean validation log likelihood for each basis, and horizontal bars show one standard deviation across blocks.

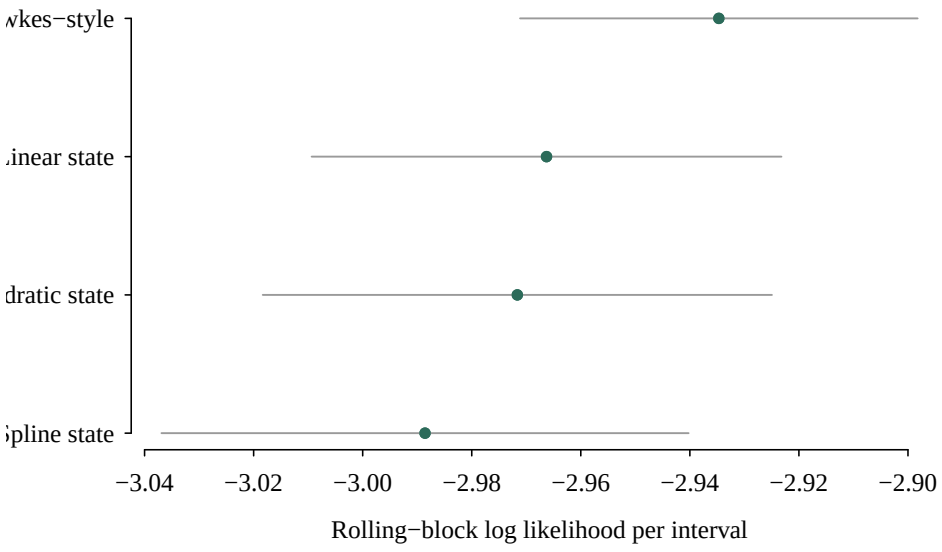


Figure 7: Held-out true and estimated demand-arrival intensities in the simulation.

The scatter plot compares the known simulated demand-arrival intensity with the estimated intensity on held-out intervals. The diagonal is perfect recovery, so deviations from the line show first-stage intensity-estimation error.

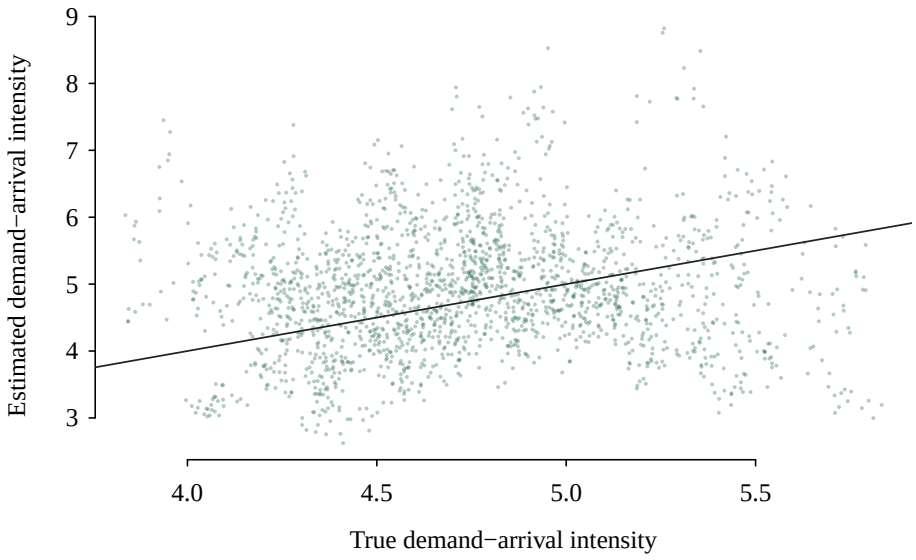



Figure 8: Generator projections are compared with reduced-form aggregate benchmarks.

The bars report held-out losses for generator-based projections and reduced-form aggregate benchmarks. Lower bars indicate better out-of-sample fit; targets differ across price adjustment and volatility.

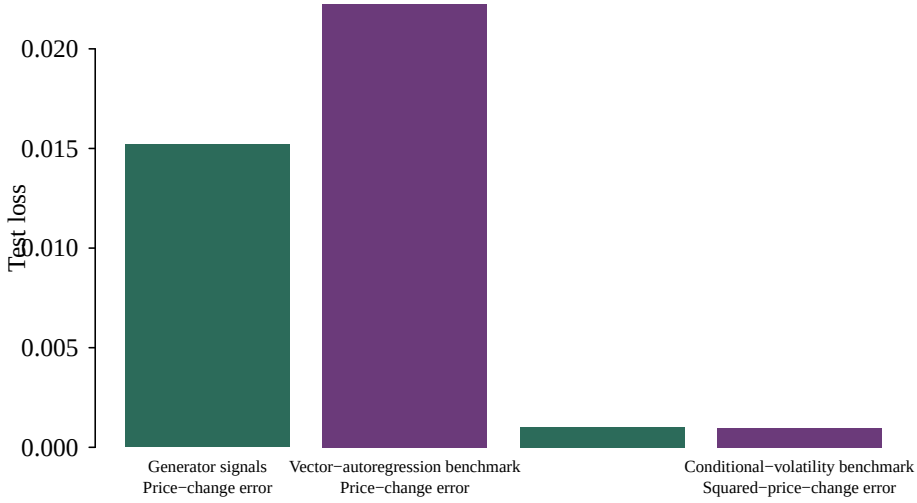


Figure 9: Aggregate price changes and one-step projections from the generator and vector-autoregression benchmark.

The time-series plot overlays realized aggregate price changes with one-step predictions from the generator projection and the vector-autoregression benchmark. The comparison shows where event-generator drift tracks the aggregate price-change path.

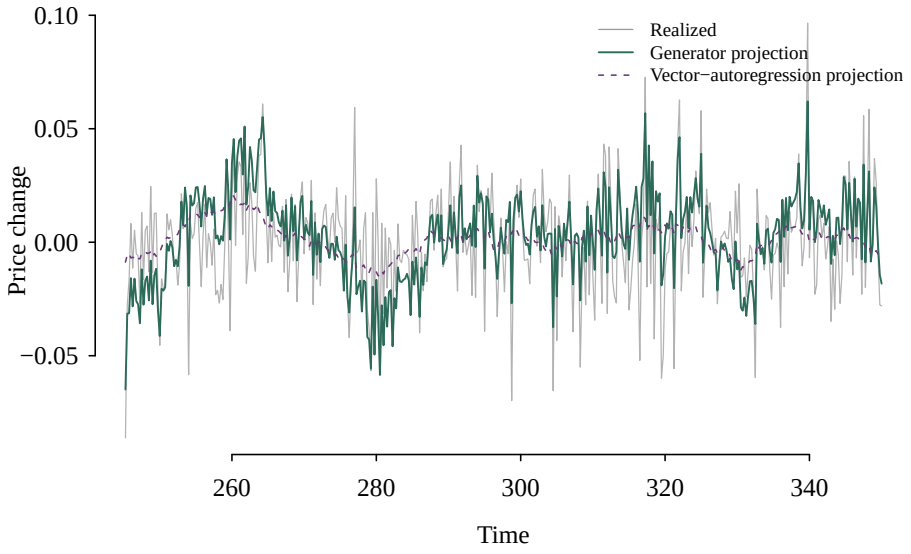
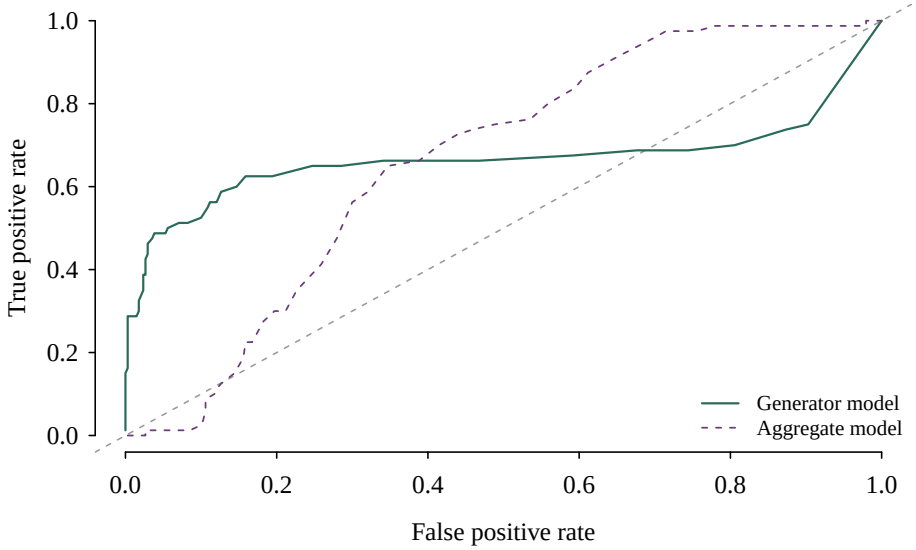



Figure 10: Rare liquidity-state classification is benchmarked against an aggregate logit.

The receiver-operating-characteristic curves compare rare liquidity-state ranking by the generator model and by the aggregate benchmark. Curves farther above the diagonal rank future low-depth or high-imbalance states more accurately.



# 9 Identification and Sieve Inference

The econometric target is the conditional adjustment law: event intensities when events are observed and transition laws when only interval movements are observed. The section separates five inferential layers. First, likelihood or transition restrictions identify the adjustment law supported by the observation regime. Second, conditional ECF restrictions identify a characteristic-function projection of that law. Third, fixed-grid and validation-selection results describe the regularized ECF estimator. Fourth, plug-in results map the selected adjustment law into observable forecast distributions. Fifth, held-out score-gap results evaluate those observable distributions against reduced-form benchmarks. This separation is necessary because a well-estimated adjustment law is not automatically a useful observable forecast.

Table 45: Inference targets in generator inversion.

The table maps each inferential layer to its target, uncertainty statement, and reported evidence. It distinguishes first-stage adjustment-law estimation from final-test forecast-score comparisons and from rare-tail diagnostics.

| Inference layer | Target | Uncertainty statement | Reported evidence |
|---|---|---|---|
| *Generator and ECF estimation* | | | |
| First-stage adjustment law | Event intensities or transition law on the visited support. | Likelihood, transition, or fixed-grid ECF consistency under the stated support and separation conditions. | Likelihood and characteristic diagnostics. |
| Adaptive ECF step | Validation-selected finite-grid characteristic projection. | Finite-grid oracle property over prespecified designs and penalties. | Selected design, penalty, and held-out characteristic loss. |
| Held-out characteristic gap | Difference between benchmark and generator characteristic losses for a fixed frequency band. | Block central limit or moving-block bootstrap for the final-test characteristic loss differential. | Characteristic gap, interval, and positive share. |
| *Observable distribution* | | | |
| Generator-implied forecast law | Observable distribution induced by the validation-selected generator and the measurement or projection law. | True observable law under a separating design and correct specification; validation-selected projection under sparse designs, regularization, or misspecification. | Quantile, interval, tail, distribution, and score-gap evidence. |
| Estimator-to-observable synthesis | Full conditional distribution of observables implied by the selected generator and measurement law. | Proposition 22 combines ECF selection, plug-in mapping, simulation approximation, and score convergence. | One inferential chain from ECF fit to observable scores. |
| Simulation approximation | Empirical forecast law computed from forward generator paths. | Monte Carlo error decreases with the number of simulated paths and is separate from first-stage estimation error. | Simulation size and distribution-score stability. |
| Held-out score gap | Difference between benchmark and generator forecast scores for a fixed target and horizon. | Block central limit or moving-block bootstrap for the final-test loss differential, with the benchmark information rule fixed before testing. | Gap estimate, interval, and positive share. |
| *Tail geometry* | | | |
| Rare-action diagnostic | Fixed-path action or smoothed rare-event functional implied by the generator. | Plug-in or bootstrap uncertainty for prespecified smooth objects; optimized rare paths need additional conditions. | Action error, action ratio, and rank correlation. |

Table 46 collects the additional symbols used in this section. The table is separate from the main

notation tables because these objects are inferential devices: they describe transition-panel densities and horizons, identified sets, asymptotic exposure, characteristic-loss uncertainty, and plug-in forecast support rather than new economic primitives.

Table 46: Additional notation for identification and inference.

The table lists the section-specific symbols used for transition-panel identification, sieve rates, held-out characteristic-loss inference, observable plug-in results, and rare-action diagnostics.

| Symbol | Name | Role in the calculation |
|---|---|---|
| $\mathcal{H}_\Delta$ | Observed horizon set | Horizons with positive probability in a transition panel. |
| $X_i^+$, $\Delta_i$ | Post-transition state and horizon | Future state and elapsed horizon in a transition-panel observation. |
| $k_0$, $k_\theta$, $Q_T(\theta)$ | Transition density and criterion | True and candidate transition densities and the population transition likelihood in (53). |
| $\Theta_{\mathrm{tr}}^*$ | Transition-equivalent generator set | Candidate generators that imply the same observed transition kernels on the visited support. |
| $d_{\mathrm{tr}}$, $K_{\mathrm{tr}}$, $b_{K_{\mathrm{tr}}}$ | Transition-law rate objects | Metric on observed-horizon transition laws, transition-class dimension, and transition approximation error. |
| $T_{\mathrm{eff}}$ | Effective event exposure | Amount of event-history information available for estimating intensity functions. |
| $K$, $a_K$ | Basis dimension and approximation error | Sieve size and approximation error in the event-intensity estimator. |
| $R_b$, $\Lambda_{\mathrm{CCF},b}$ | Band-loss matrix and long-run variance | Objects used to compute and infer held-out characteristic-loss gaps in frequency band $b$. |
| $\hat{\alpha} = (\hat{d}, \hat{\rho})$ | Selected ECF rule | Validation-selected characteristic design and regularization value used in the final generator estimate. |
| $\theta^*$ | Selected population target | Population target associated with the selected ECF design and regularization rule. |
| $\mathcal{X}_0$ | Visited forecast support | State support on which the observable plug-in distribution is evaluated. |
| $F_{Y,\theta^*}$ | Target observable law | Conditional observable distribution implied by the selected population generator target and measurement law. |
| $\mathcal{B}_{\mathrm{bench}}$, $\hat{b}$, $U(b)$, $\hat{U}(b)$ | Benchmark-selection objects | Prespecified benchmark library, selected benchmark, population validation score, and sample validation score. |
| $F_{G,i}^*$, $F_{B,i}^*$, $d_i^*$, $\Lambda_S$ | Score-gap limit objects | Limiting generator and benchmark forecast laws, limiting loss differential, and long-run variance for held-out score-gap inference. |
| $h_{\mathrm{act}}$, $I_{\mathrm{act}}$ | Action horizon and fixed-path action | Prespecified horizon and action value used in the rare-event path diagnostic. |

Three targets should be kept separate throughout the section. With event histories and support separation, the structural target is the local adjustment generator on the exposed state support. With finite ECF grids, regularization, and validation, the statistical target is the selected characteristic projection, even when no structural generator is point-identified. After a measurement law is fixed, the evaluation target is the induced observable forecast distribution and its score gap against prespecified reduced-form benchmarks. Correct specification and separating restrictions make these targets coincide. Sparse grids, regularization, transition-only data, or misspecification deliberately keep them apart. The rest of the section proves the links in this hierarchy rather than treating generator fit and observable forecast value as the same claim.

The propositions should be read as a composition argument. Identification results state what

the observation regime can recover before any forecast score is computed. The ECF results then define the finite-grid or validation-selected target actually estimated in finite samples. The mapping results translate that selected adjustment law into a conditional distribution for observables. The final score-gap results evaluate that distribution on untouched test data. This ordering is part of the estimand: changing the order, for example by choosing frequencies after seeing final forecast scores, would change the statistical target and would no longer support the same claim.

In the event-history case, let $N_r(t)$ be the counting process for event type $r$, let $\mathcal{F}_t$ denote the event history and observed state, and suppose

$$\mathbb{E}[\mathrm{d}N_r(t) \mid \mathcal{F}_{t-}] = \lambda_{r,0}(X_{t-})\mathrm{d}t. \tag{50}$$

The statistical object is $\lambda_0 = \{\lambda_{r,0} : r \in \mathcal{R}\}$, not only a discretely sampled aggregate process. Under this formulation, observed arrivals, cancellations, removals, revisions, transitions, or other adjustment events identify the generator through the compensator of the counting process. For lower-frequency panels, the analogous object is the conditional transition law of the increment given the observed state. This is the same likelihood logic used in event-history and counting-process models (Andersen and Gill, 1982), with the conditional ECF criterion supplying the distributional comparison when a full transition density is inconvenient or deliberately kept semiparametric.

**Assumption 1** (Observed event support)**.** *For each event type $r$, the process $N_r(t) - \int_0^t \lambda_{r,0}(X_{s-})\mathrm{d}s$ is an $\mathcal{F}_t$-martingale. The intensity is predictable and finite on the support reached by the data. If two candidate intensities $\lambda_r$ and $\tilde{\lambda}_r$ differ on a set with positive integrated exposure,*

$$\mathbb{E}\int_0^T \mathbb{1}\{\lambda_r(X_{t-}) \neq \tilde{\lambda}_r(X_{t-})\}\mathrm{d}t > 0,$$

*then the two candidates are treated as different economic generators.*

For a candidate generator $\lambda = \{\lambda_r\}$, the population criterion is the expected log likelihood

$$Q(\lambda) = \sum_{r\in\mathcal{R}} \mathbb{E}\left[\int \log \lambda_r(X_{t-})\mathrm{d}N_r(t) - \int \lambda_r(X_{t-})\mathrm{d}t\right]. \tag{51}$$

Using the true compensator, the difference between a candidate and the truth is

$$Q(\lambda) - Q(\lambda_0) = -\sum_{r\in\mathcal{R}} \mathbb{E} \int \left[\lambda_{r,0}(X_{t-}) \log \frac{\lambda_{r,0}(X_{t-})}{\lambda_r(X_{t-})} - \lambda_{r,0}(X_{t-}) + \lambda_r(X_{t-})\right] \mathrm{d}t. \tag{52}$$

The bracketed term is the Poisson Kullback–Leibler divergence. Hence the population criterion is uniquely maximized by the true generator whenever two candidate intensities that differ on a set visited with positive exposure are treated as different models. This is the basic identification condition.

**Proposition 5** (Likelihood identification of the generator)**.** *Under Assumption 1, $Q(\lambda) \le Q(\lambda_0)$. Equality holds only for generators that agree with $\lambda_0$ on the event support, up to observational equivalence.*

*Proof.* Equation (52) writes the population likelihood loss as the negative of an integrated Poisson Kullback–Leibler divergence. The divergence is nonnegative pointwise and equals zero only when $\lambda_r(X_{t-}) = \lambda_{r,0}(X_{t-})$. Assumption 1 rules out differences on sets with positive exposure, so the population criterion identifies the generator on the observed support. □

For transition panels, the local events may not be observed. Let the data be triples $(X_i, X_i^+, \Delta_i)$, where $X_i^+$ is the state observed after horizon $\Delta_i$. A candidate generator implies a transition kernel $K_{\theta,\Delta_i}(\mathrm{d}x' \mid X_i)$. If this kernel has density $k_\theta(x' \mid x, \Delta)$ with respect to a common dominating measure, the population transition criterion is

$$Q_T(\theta) = \mathbb{E}[\log k_\theta(X_i^+ \mid X_i, \Delta_i)]. \tag{53}$$

**Assumption 2** (Observed transition support)**.** *The true conditional transition law has density $k_0(x' \mid x, \Delta)$ on the visited support. Candidate kernels are finite and correctly normalized on that support. If two candidates imply different transition laws on a set with positive probability,*

$$\Pr\{K_{\theta,\Delta_i}(\cdot \mid X_i) \neq K_{\tilde{\theta},\Delta_i}(\cdot \mid X_i)\} > 0,$$

*then the two candidates are treated as different transition laws.*

**Proposition 6** (Transition-law identification)**.** *Under Assumption 2, the transition likelihood is*

*maximized at the true transition law:*

$$Q_T(\theta) \le Q_T(\theta_0).$$

*Equality holds only for candidates that imply the same conditional transition law as the truth on the visited support.*

*Proof.* The criterion difference satisfies

$$Q_T(\theta) - Q_T(\theta_0) = -\mathbb{E}\left[\int \log \frac{k_0(x' \mid X_i, \Delta_i)}{k_\theta(x' \mid X_i, \Delta_i)} K_{0,\Delta_i}(\mathrm{d}x' \mid X_i)\right],$$

which is the negative conditional Kullback–Leibler divergence integrated over the visited state and horizon support. The divergence is nonnegative and equals zero only when the candidate and true transition kernels agree almost surely. □

The previous proposition identifies the observed transition law, not automatically the continuous-time generator behind that law. Let $\mathcal{H}_\Delta$ denote the set of horizons observed with positive probability. The transition-equivalent generator set is

$$\Theta_{\mathrm{tr}}^* = \{\theta \in \Theta : K_{\theta,\Delta}(\cdot \mid x) = K_{0,\Delta}(\cdot \mid x) \text{ for all } \Delta \in \mathcal{H}_\Delta \text{ and visited } x\}.$$

**Assumption 3** (Transition-to-generator injectivity)**.** *If two candidate generators imply the same transition kernels on the visited state support for every horizon in $\mathcal{H}_\Delta$, then they are observationally equivalent for the generator features used in the paper:*

$$K_{\theta,\Delta}(\cdot \mid x) = K_{\tilde{\theta},\Delta}(\cdot \mid x) \quad \textit{for all } \Delta \in \mathcal{H}_\Delta \textit{ and visited } x \quad \Rightarrow \quad \theta \sim \tilde{\theta}.$$

*If the paper reports counterfactual horizons or path objects outside $\mathcal{H}_\Delta$, this equivalence must be strengthened to equality of the corresponding semigroup or path functional.*

**Proposition 7** (Generator identification from transition panels)**.** *Under Assumption 2, transition data identify $\Theta_{\mathrm{tr}}^*$. Under Assumption 3, this set collapses to the generator parameter up to observational equivalence. Without this injectivity condition, observable forecasts are identified only for functionals that are invariant over $\Theta_{\mathrm{tr}}^*$, or for horizons and mappings fixed directly by the observed*

*transition law.*

*Proof.* Proposition 6 shows that all population maximizers imply the true transition kernels on the visited state and horizon support, so the identified set is $\Theta_{\mathrm{tr}}^*$. Assumption 3 maps equality of those kernels into equality of the generator features used for estimation, forecasting, and testing. If the assumption is not imposed, any two elements of $\Theta_{\mathrm{tr}}^*$ are observationally indistinguishable in the transition panel, and only functionals that agree across that set are identified. □

For transition panels the first-stage rate is a rate for the observed-horizon transition law. This is enough for observable forecasting at the same horizons, but it is weaker than an event-history generator rate unless the injectivity or semigroup conditions above are imposed. Let $d_{\mathrm{tr}}$ be a metric on transition laws over the visited state and horizon support, for example

$$d_{\mathrm{tr}}^2(\theta, \theta_0) = \mathbb{E}\left[\|K_{\theta,\Delta_i}(\cdot \mid X_i) - K_{0,\Delta_i}(\cdot \mid X_i)\|_{\mathrm{BL}}^2\right],$$

where $\|\cdot\|_{\mathrm{BL}}$ is the bounded-Lipschitz distance between probability laws. Let $K_{\mathrm{tr}}$ denote the dimension of the transition-law approximation used for the anchor or fixed-grid ECF fit, and let $b_{K_{\mathrm{tr}}}$ be the approximation error in $d_{\mathrm{tr}}$.

**Assumption 4** (Transition-panel rate regularity)**.** *The transition observations $(X_i, X_i^+, \Delta_i)$ are stationary and absolutely regular with mixing coefficients strong enough for a block central limit theorem and uniform laws of large numbers for the transition criterion. The candidate transition class contains an approximation with error $b_{K_{\mathrm{tr}}}$ in $d_{\mathrm{tr}}$. The population transition likelihood or fixed-grid transition ECF criterion has local quadratic curvature in $d_{\mathrm{tr}}$ around its pseudo-true transition law after removing observationally equivalent directions. The stochastic part of the $K_{\mathrm{tr}}$-dimensional criterion is $O_p\{(K_{\mathrm{tr}} \log n/n)^{1/2}\}$, and $K_{\mathrm{tr}} \log n/n \to 0$.*

**Proposition 8** (Transition-panel estimation rate)**.** *Let $\hat{\theta}_{\mathrm{tr}}$ be an approximate optimizer of the transition likelihood, a prespecified transition fit, or a fixed-grid transition ECF criterion over the $K_{\mathrm{tr}}$-dimensional transition class. Under Assumptions 2 and 4,*

$$d_{\mathrm{tr}}(\hat{\theta}_{\mathrm{tr}}, \theta_0) = O_p\left(\sqrt{\frac{K_{\mathrm{tr}} \log n}{n}}\right) + O(b_{K_{\mathrm{tr}}}).$$

*Under Assumption 3, any Lipschitz generator feature or observable forecast functional inherits this*

*rate on the visited support. Without the injectivity condition, the rate applies to the observed-horizon transition law and to observable forecast functionals determined by that transition law, not to a unique continuous-time generator.*

*Proof.* The approximation term is the distance from the true transition law to the best element of the $K_{\mathrm{tr}}$-dimensional transition class. Local quadratic curvature converts excess population criterion value into squared $d_{\mathrm{tr}}$-distance. The uniform stochastic bound controls the empirical criterion over the transition class. Combining the approximation and stochastic terms gives the displayed rate by the standard sieve $M$-estimation argument. The final statement follows by the Lipschitz continuous mapping argument for generator or observable functionals, with Assumption 3 determining whether the transition law identifies a unique generator feature. □

Table 47 summarizes the claim discipline. Event histories are the cleanest route to a local generator because the compensator directly identifies event intensities on exposed support. Transition panels first identify transition laws. The paper can call the fitted object a generator only when a semigroup restriction, multiple horizons, or an injective parameterization links the observed transition kernels to the underlying generator. A finite ECF design adds a further distinction: it identifies a finite-grid distributional projection unless the selected frequencies and instruments separate the conditional law on the visited support.

Table 47: Identified objects by data regime.

The table states what is identified from event histories, transition likelihoods, finite ECF designs, and final observable forecast scores. It separates local generator identification from transition-law and finite-grid projection claims.

| Evidence source | Identified object | Additional condition for a generator claim | Claim scope |
|---|---|---|---|
| Event histories with exposure | Local adjustment intensities on the exposed state support. | Counting-process compensator and support separation for the observed event types. | Local generator. |
| Transition likelihood at observed horizons | Conditional transition law at the visited states and horizons. | Semigroup restrictions, multiple horizons, or an injective transition-to-generator parameterization. | Generator only under extra structure. |
| Conditional ECF with a finite design | Best finite-grid distributional projection of the increment law. | Characteristic separation and an injective generator parameterization on the visited support. | Projection unless separating. |
| Observable forecast scores | Held-out value of the generator-implied observable distribution. | Measurement or projection law fixed before final testing, with benchmarks chosen before the test block. | Target, horizon, and score. |

Likelihood is not the only identification route. The ECF criterion identifies the conditional distribution of increments through the characteristic functions generated by candidate adjustment laws.

This matters for transition data when the transition density is unavailable, expensive to evaluate, or deliberately replaced by simulation, and for semiparametric settings where the researcher wants to target distributional implications without specifying a full density.

**Assumption 5** (Characteristic separation)**.** *Let $\phi_0(u \mid x, \Delta)$ denote the true conditional characteristic function of the observed increment over horizon $\Delta$, and let $\phi_\theta(u \mid x, \Delta)$ be the characteristic function implied by a candidate generator. The instrument class is rich enough that*

$$\mathbb{E}[z(X_t)\{\phi_0(u \mid X_t, \Delta) - \phi_\theta(u \mid X_t, \Delta)\}] = 0$$

*for all instruments $z$ in the class and all frequencies $u$ in a neighborhood of the origin only if $\phi_\theta(u \mid X_t, \Delta) = \phi_0(u \mid X_t, \Delta)$ almost surely for those frequencies. The generator parameterization is identified up to observational equivalence by this conditional characteristic function on the visited state support.*

**Proposition 9** (Conditional ECF identification)**.** *Under Assumption 5, the population conditional ECF criterion is minimized by generators that reproduce the true conditional transition law of the observed increment on the visited support. If the generator parameterization is injective on that support, the population minimizer is the true generator parameter.*

*Proof.* If the population moments are zero for all instruments and frequencies in the separating class, Assumption 5 gives equality of the conditional characteristic functions almost surely. Characteristic functions determine probability laws, so the conditional transition law of the increment is the same under the candidate and the truth. Injectivity of the generator parameterization then gives equality of the generator parameter up to observational equivalence. □

The implemented estimator uses finite frequency and instrument grids. These grids identify the best finite-dimensional distributional projection, not the entire conditional law. This distinction is important. The grid is an empirical design choice, and the validation step asks whether that finite set of characteristic restrictions improves held-out distributional fit relative to the likelihood or transition anchor.

Formally, for fixed frequency and instrument sets $(\mathcal{U}, \mathcal{Z})$, let

$$m_{\mathcal{U}\mathcal{Z}}(\theta) = \mathbb{E}\left[z_j(X_i)\left\{\exp(\mathrm{i}u_m^\top \Delta X_i) - \phi_\theta(u_m \mid X_i, \Delta_i)\right\}\right]_{m,j}$$

denote the stacked population characteristic moments, including real and imaginary parts. For a population weighting matrix $W$, the finite-grid ECF target is

$$\Theta^*_{\mathcal{U}\mathcal{Z}} = \arg\min_{\theta\in\Theta} m_{\mathcal{U}\mathcal{Z}}(\theta)' W m_{\mathcal{U}\mathcal{Z}}(\theta).$$

If the model is correctly specified and the grid is separating on the visited support, this set contains the true generator parameter. If the model is misspecified or the grid is deliberately sparse, $\Theta^*_{\mathcal{U}\mathcal{Z}}$ is the pseudo-true generator projection that best matches the selected characteristic restrictions. For a population anchor $\theta_{\mathrm{A}}$ and penalty $\rho$, the corresponding regularized target is

$$\theta^*_\rho \in \arg\min_{\theta\in\Theta}\left\{m_{\mathcal{U}\mathcal{Z}}(\theta)' W m_{\mathcal{U}\mathcal{Z}}(\theta) + \rho(\theta - \theta_{\mathrm{A}})'\Omega(\theta - \theta_{\mathrm{A}})\right\}.$$

The validation rule then selects among these population targets by their out-of-sample characteristic loss $V(\rho)$. The selected object is therefore a validation-chosen distributional projection. It equals the structural generator only under the stronger correct-specification and separation conditions; otherwise it is still an interpretable forecast object because its performance is judged on held-out characteristic and observable distribution scores.

For a fixed frequency and instrument grid, the regularized ECF estimator is an ordinary finite-dimensional minimum-distance estimator. This observation is useful because it separates two issues. The first is standard estimation error for a prespecified grid and penalty. The second is empirical design: how the grid, penalty, basis, and final test block are chosen.

Let $\hat{m}_n(\theta)$ be the sample analogue of $m_{\mathcal{U}\mathcal{Z}}(\theta)$, and define the sample criterion

$$\hat{Q}_{\rho,n}(\theta) = \hat{m}_n(\theta)' W_n \hat{m}_n(\theta) + \rho(\theta - \hat{\theta}_{\mathrm{A}})'\Omega_n(\theta - \hat{\theta}_{\mathrm{A}}).$$

The corresponding population criterion is $Q_\rho(\theta)$, obtained by replacing $\hat{m}_n$, $W_n$, $\Omega_n$, and $\hat{\theta}_{\mathrm{A}}$ by their probability limits. The next assumption records the regularity needed for the fixed-grid target.

**Assumption 6** (Fixed-grid ECF regularity)**.** *For each fixed $\rho$, the population criterion $Q_\rho$ has a*

*unique minimizer* $\theta_\rho^*$ *in the interior of* $\Theta$*. The sample moments satisfy* $\sup_{\theta\in\Theta} \|\hat{m}_n(\theta) - m_{\mathcal{UZ}}(\theta)\| \to_p 0$*, and* $W_n \to_p W$*,* $\Omega_n \to_p \Omega$*, and* $\hat{\theta}_{\mathrm{A}} \to_p \theta_{\mathrm{A}}$*. In a neighborhood of* $\theta_\rho^*$*, the sample criterion is twice continuously differentiable with a nonsingular population Hessian* $H_\rho = \nabla^2 Q_\rho(\theta_\rho^*)$*, and the sample score obeys*

$$\sqrt{n}\,\nabla\hat{Q}_{\rho,n}(\theta_\rho^*) \Rightarrow N(0, \Sigma_\rho).$$

**Proposition 10** (Fixed-grid regularized ECF estimation)**.** *Under Assumption 6, any approximate minimizer* $\hat{\theta}_\rho$ *of* $\hat{Q}_{\rho,n}$ *satisfies*

$$\hat{\theta}_\rho \to_p \theta_\rho^*.$$

*If the local differentiability and score central limit condition in Assumption 6 also hold, then*

$$\sqrt{n}(\hat{\theta}_\rho - \theta_\rho^*) = -H_\rho^{-1}\sqrt{n}\,\nabla\hat{Q}_{\rho,n}(\theta_\rho^*) + o_p(1),$$

*and hence*

$$\sqrt{n}(\hat{\theta}_\rho - \theta_\rho^*) \Rightarrow N(0, H_\rho^{-1}\Sigma_\rho H_\rho^{-1}).$$

*Proof.* Uniform convergence of the sample criterion to $Q_\rho$, together with uniqueness of $\theta_\rho^*$, gives consistency by the argmin theorem. For the local law, expand the first-order condition around $\theta_\rho^*$. The Hessian converges to the nonsingular matrix $H_\rho$, and the sample score has the stated central limit law. Solving the linear expansion gives the displayed asymptotic representation and covariance formula. □

Proposition 10 is intentionally a fixed-grid statement. It says what is estimated once the frequencies, instruments, weighting rule, anchor, and penalty are fixed. It does not by itself justify searching over many grids or penalties on the final test block. That is why the implemented procedure uses a separate validation block for ECF-design and regularization choice and reserves the final block for reporting characteristic and observable forecast scores.

The adaptive target is therefore a selected rule, not a post hoc best-looking generator. Let $\alpha = (d, \rho)$ denote a characteristic design and penalty pair, where $d$ fixes the frequency grid, instruments, block length, and weighting rule. Let $\theta_\alpha^*$ be the regularized population ECF target for that pair. The

population validation criterion is

$$V(\alpha) = L_{\mathrm{CCF},\mathcal{V}}(\theta_\alpha^*, d) + \xi p_{\psi,d}/n_{\mathcal{V}},$$

where $L_{\mathrm{CCF},\mathcal{V}}$ is the population counterpart of the validation characteristic loss and $p_{\psi,d}$ is the number of stacked real moments in design $d$. The selected population object is any minimizer of $V(\alpha)$ over the prespecified finite library. If the minimizer is unique and separated, the adaptive estimator targets that design and penalty. If several designs are tied, the target is the forecast law generated by the prespecified selection rule. This convention keeps regularization and adaptivity inside the estimator, while the final test block remains reserved for evidence.

**Proposition 11** (Validation selection of the adaptive regularized ECF step)**.** *Let $\mathcal{D}$ be a finite set of prespecified ECF designs and let $\mathcal{P}$ be a finite set of regularization choices, including the likelihood or transition anchor. For each $\alpha = (d, \rho) \in \mathcal{D} \times \mathcal{P}$, let $\hat{\theta}_\alpha$ minimize the corresponding regularized ECF criterion on the estimation sample, and let $\hat{V}(\alpha)$ be the validation characteristic loss, including any prespecified complexity price, computed on a separate block. If*

$$\sup_{\alpha \in \mathcal{D}\times\mathcal{P}} |\hat{V}(\alpha) - V(\alpha)| \to_p 0,$$

*then the selected design and penalty $\hat{\alpha} = (\hat{d}, \hat{\rho})$ satisfy the finite-grid oracle inequality*

$$V(\hat{\alpha}) \leq \min_{\alpha \in \mathcal{D}\times\mathcal{P}} V(\alpha) + o_p(1).$$

*If the population minimizer set is separated from the rest of the grid, the selected pair lies in that minimizer set with probability approaching one. In particular, if the anchor under a low-dimensional design is the separated population minimizer, the procedure keeps that conservative specification with probability approaching one. If a richer ECF design improves validation loss by a fixed gap after the complexity price, the procedure selects a richer non-anchor ECF fit with probability approaching one.*

*Proof.* Let $\alpha^* \in \arg\min_{\alpha \in \mathcal{D}\times\mathcal{P}} V(\alpha)$. Since $\hat{\alpha}$ minimizes $\hat{V}$,

$$V(\hat{\alpha}) \leq \hat{V}(\hat{\alpha}) + o_p(1) \leq \hat{V}(\alpha^*) + o_p(1) \leq V(\alpha^*) + o_p(1).$$

This gives the oracle inequality. If the minimizer set is separated, uniform convergence over the finite grid rules out selection of any design–penalty pair outside that set with probability approaching one. The conservative and richer-design statements are the two empirically relevant separated cases. □

The final characteristic comparison is a test-block object, not another tuning criterion. Let $\hat{g}_{\mathcal{T},a,b}$ denote the stacked held-out characteristic moments for candidate model $a$ and frequency band $b$, using the same instruments as in (18). Let $R_b$ be the fixed averaging matrix that turns $\hat{g}_{\mathcal{T},a,b}$ into the reported band loss, so that $L_{\mathrm{CCF}}(a,b) = \hat{g}'_{\mathcal{T},a,b} R_b \hat{g}_{\mathcal{T},a,b}$. For a generator $G$ and benchmark $B$, define the held-out characteristic-loss gap

$$\hat{\Delta}_{\mathrm{CCF},b}(G,B) = L_{\mathrm{CCF}}(B,b) - L_{\mathrm{CCF}}(G,b).$$

Positive values mean that the generator has smaller held-out characteristic residuals than the benchmark in frequency band $b$.

**Assumption 7** (Held-out characteristic-loss regularity)**.** *The generator, benchmark, characteristic design, regularization, and weighting rule are fixed before the final test block is evaluated. The stacked test-block moment vector*

$$\sqrt{n_{\mathcal{T}}}\left[\begin{pmatrix}\hat{g}_{\mathcal{T},G,b}\\ \hat{g}_{\mathcal{T},B,b}\end{pmatrix} - \begin{pmatrix}g^*_{G,b}\\ g^*_{B,b}\end{pmatrix}\right]$$

*satisfies a block central limit theorem with long-run covariance matrix $\Sigma_{\mathrm{CCF},b}$. The quadratic loss-gap map is first-order nondegenerate at $(g^*_{G,b}, g^*_{B,b})$, or the reported uncertainty is based directly on the bootstrap distribution of the quadratic loss gap.*

**Proposition 12** (Held-out characteristic-loss gap inference)**.** *Under Assumption 7,*

$$\hat{\Delta}_{\mathrm{CCF},b}(G,B) \to_p \Delta_{\mathrm{CCF},b}(G,B) = g^{*\prime}_{B,b} R_b g^*_{B,b} - g^{*\prime}_{G,b} R_b g^*_{G,b}.$$

*In the nondegenerate first-order case,*

$$\sqrt{n_{\mathcal{T}}}\{\hat{\Delta}_{\mathrm{CCF},b}(G,B) - \Delta_{\mathrm{CCF},b}(G,B)\} \Rightarrow N(0, \Lambda_{\mathrm{CCF},b})$$

*for the delta-method long-run variance $\Lambda_{\mathrm{CCF},b}$. If a moving-block bootstrap consistently estimates the law of the stacked characteristic moments, then the bootstrap distribution of the recomputed*

*characteristic-loss gap consistently estimates the sampling law of the reported gap.*

*Proof.* Consistency follows from convergence of the stacked characteristic moments and continuity of the quadratic loss map. When the derivative of the quadratic loss-gap map is nonzero at the population moment vector, the delta method applied to the block central limit theorem gives the normal limit. Recomputing the quadratic loss gap inside each moving-block bootstrap draw applies the same continuous mapping to the bootstrap moment law, which yields the stated bootstrap approximation. □

The sieve estimator restricts the log intensity to a finite basis:

$$\lambda_{r,K}(x;\theta_r) = e_r(x)\exp\{\theta_r^\top \phi_K(x)\}. \tag{54}$$

The estimator maximizes the sample likelihood over $\theta_r$ for each event type. This formulation is deliberately close to standard series estimation (Newey, 1997; Chen, 2007): the basis dimension $K$ trades approximation bias against sampling variance.

**Assumption 8** (High-level sieve regularity)**.** *For each event type, the sieve class contains an approximation $\lambda^*_{r,K}$ with error*

$$\|\lambda^*_{r,K} - \lambda_{r,0}\| \le a_K, \qquad a_K \to 0.$$

*The empirical likelihood is locally well behaved around $\lambda_0$: the information matrix on the visited support is nonsingular after removing unidentified directions, the effective exposure $T_{\text{eff}}$ grows, and the stochastic part of the score over the $K$-dimensional sieve is $O_p(\sqrt{K/T_{\text{eff}}})$.*

If the log intensity is smooth and the chosen basis approximates it well, the approximation error decreases with $K$. At the same time, the estimation error grows with the number of basis coefficients relative to effective event exposure. Under Assumption 8, the rate implication is

$$\|\hat{\lambda}_K - \lambda_0\| = O_p\left(\sqrt{\frac{K}{T_{\text{eff}}}}\right) + O\left(a_K\right), \tag{55}$$

where $a_K$ is the basis approximation error and $T_{\text{eff}}$ denotes the amount of event exposure. Equation (55) is not a primitive assumption that makes the empirical problem disappear. It states the burden of proof: an empirical paper must control both sampling error and basis approximation error.

A more explicit route to (55) is useful for the empirical paper. Define the generator norm

$$\|\lambda - \lambda_0\|_T^2 = \sum_{r \in \mathcal{R}} \frac{1}{T_{\text{eff}}} \mathbb{E} \int_0^T \{\lambda_r(X_{t-}) - \lambda_{r,0}(X_{t-})\}^2 \mathrm{d}t. \tag{56}$$

This norm measures errors on the state support that is actually exposed to event or adjustment risk. It is the relevant norm for economic projections because unvisited states cannot be learned from the observed history.

**Assumption 9** (Primitive rate conditions)**.** *The observed state process is stationary and absolutely regular, with mixing coefficients that decay fast enough for a block central limit theorem for the likelihood score. For each event type, the true intensity is bounded away from zero and infinity on the visited support. The basis satisfies* $\sup_x \|\phi_K(x)\| \lesssim \sqrt{K}$*, the population information matrix has eigenvalues bounded away from zero and infinity after removing observationally equivalent directions, and there exists a sieve approximation with error* $a_K$ *in the norm* (56)*. The dimension grows slowly enough that* $K \log T_{\text{eff}}/T_{\text{eff}} \to 0$*.*

**Proposition 13** (Sieve rate for event-generator estimation)**.** *Under Assumptions 1 and 9, the point-process sieve likelihood estimator satisfies*

$$\|\hat{\lambda}_K - \lambda_0\|_T = O_p\left(\sqrt{\frac{K \log T_{\text{eff}}}{T_{\text{eff}}}}\right) + O(a_K).$$

*Proof.* The likelihood identification argument gives population curvature in the generator norm around $\lambda_0$. The mixing and bounded-basis conditions control the empirical score uniformly over the $K$-dimensional sieve, producing the stochastic term. The approximation term is the distance from the true log intensity to the sieve. This is the counting-process analogue of the usual series-estimation bias–variance decomposition in Newey (1997) and Chen (2007). □

**Proposition 14** (Prespecified basis selection)**.** *Let* $\mathcal{B}_{\text{basis}}$ *be a finite set of candidate bases, and let* $\hat{b}$ *maximize a validation likelihood computed on a block of event histories not used for the final test comparison. If the population validation criterion has a unique maximizer* $b_0$ *separated from all other bases by more than the uniform validation error, then* $\Pr(\hat{b} = b_0) \to 1$*. If two bases are tied at the population criterion, the estimand is the performance of the prespecified selection rule rather than the performance of a single fixed basis.*

*Proof.* Uniform convergence of the validation likelihood over the finite set $\mathcal{B}_{\text{basis}}$ implies that the sample maximizer equals the separated population maximizer with probability approaching one. When the gap is local or zero, model selection remains part of the estimator. Reporting the selected-rule target avoids conditioning on a basis as if it had been fixed ex ante. □

The generator projections inherit this uncertainty. For excess demand, the drift and variance functionals are

$$A_X(\lambda)(x) = \lambda_{D^+}(x) + \lambda_{S^-}(x) - \lambda_{D^-}(x) - \lambda_{S^+}(x), \tag{57}$$

$$B_X(\lambda)(x) = \lambda_{D^+}(x) + \lambda_{S^-}(x) + \lambda_{D^-}(x) + \lambda_{S^+}(x). \tag{58}$$

These are smooth functionals of the event intensities. Standard delta-method or bootstrap arguments can therefore be used for projections such as drift curves, volatility curves, and rare-event scores, provided the first-stage generator estimator is regular enough. The goal is not only to estimate event intensities, but also to quantify uncertainty for the aggregate econometric objects derived from them.

The conditional characteristic function is another smooth projection over any bounded frequency set. For fixed $u$ and $\Delta$,

$$\phi_\lambda(u \mid x, \Delta) = \exp\{\Delta H_\lambda(x, \mathrm{i}u)\}$$

is continuously differentiable in the event intensities whenever those intensities are finite on the visited support. Consequently, the first-stage generator rate transfers to characteristic-function diagnostics at fixed frequencies. This is the distributional analogue of transferring intensity error to drift or variance curves: estimating the event generator identifies not only low-order moments but the local conditional characteristic law induced by the jump process.

A feasible resampling design has to begin before the projection is formed. The bootstrap therefore resamples contiguous event blocks from the estimation window, reselects the state basis by validation likelihood, refits the event-intensity models, and then recomputes the aggregate drift, variance, price-adjustment, volatility, and rare-event gaps on the fixed test window. Under a fixed selected basis, or under the separated selection case in Proposition 14, the target is the distribution of the selected generator projection.

**Proposition 15** (Bootstrap for smooth projections)**.** *Let $G(\lambda)$ be a finite-dimensional projection,*

*such as a drift curve, variance curve, or smoothed forecast-loss gap, that is continuously differentiable at $\lambda_0$. Suppose the selected basis is fixed with probability approaching one, the block length $L$ satisfies $L \to \infty$ and $L/T_{\text{eff}} \to 0$, and the moving-block bootstrap consistently estimates the score distribution under Assumption 9. Then, conditionally on the data,*

$$\sqrt{T_{\text{eff}}}\{G(\hat{\lambda}^*_{\hat{b}}) - G(\hat{\lambda}_{\hat{b}})\} \Rightarrow_* \mathcal{L}\left[\sqrt{T_{\text{eff}}}\{G(\hat{\lambda}_{\hat{b}}) - G(\lambda_0)\}\right],$$

*where $\Rightarrow_*$ denotes weak convergence of the bootstrap law in probability.*

*Proof.* Proposition 14 reduces separated selection to a fixed-basis problem with probability approaching one. The block bootstrap reproduces the first-stage score law for dependent event histories. Continuous differentiability of $G$ then transfers the bootstrap approximation from the generator coefficients to the projection by the functional delta method. □

The fixed-path rare-action diagnostic in Section 8 is also a generator projection after the path has been chosen. For displayed depth, write

$$a_\lambda(x) = \lambda_{D^+}(x) + \lambda_{S^+}(x), \qquad b_\lambda(x) = \lambda_{D^-}(x) + \lambda_{S^-}(x).$$

For a prespecified absolutely continuous path $\Gamma = \{(q_s, x_s) : 0 \le s \le h_{\text{act}}\}$, with velocity $v_s = \dot{q}_s$, define the plug-in action

$$I_{\text{act}}(\Gamma; \lambda) = \int_0^{h_{\text{act}}} L\{v_s; a_\lambda(x_s), b_\lambda(x_s)\}\mathrm{d}s, \tag{59}$$

where

$$L(v; a, b) = \sup_{\varpi}\{\varpi v - a(e^{\varpi} - 1) - b(e^{-\varpi} - 1)\}.$$

**Proposition 16** (Fixed-path action plug-in)**.** *Suppose that, in a neighborhood of $\lambda_0$, the depth-increasing and depth-decreasing rates along $\Gamma$ satisfy*

$$0 < c \le a_\lambda(x_s), b_\lambda(x_s) \le C < \infty \quad \textit{and} \quad |v_s| \le V$$

*for all $s \in [0, h_{\text{act}}]$. Then $I_{\text{act}}(\Gamma; \lambda)$ is locally Lipschitz and continuously differentiable in the rate*

*paths $a_\lambda(x_s)$ and $b_\lambda(x_s)$. In particular, for a constant $C_\Gamma$,*

$$|I_{\mathrm{act}}(\Gamma;\hat{\lambda}) - I_{\mathrm{act}}(\Gamma;\lambda_0)| \le C_\Gamma \sup_{0\le s\le h_{\mathrm{act}}} \left[|a_{\hat{\lambda}}(x_s) - a_{\lambda_0}(x_s)| + |b_{\hat{\lambda}}(x_s) - b_{\lambda_0}(x_s)|\right].$$

*If the first-stage estimator controls this path norm at rate $r_T = \sqrt{K \log T_{\mathrm{eff}}/T_{\mathrm{eff}}} + a_K$, then*

$$I_{\mathrm{act}}(\Gamma;\hat{\lambda}) - I_{\mathrm{act}}(\Gamma;\lambda_0) = O_p(r_T).$$

*Proof.* Let $y = e^p$. The maximizer solves

$$ay^2 - vy - b = 0, \qquad y(v,a,b) = \frac{v + \sqrt{v^2 + 4ab}}{2a}.$$

On the compact set $a, b \in [c, C]$ and $|v| \le V$, this optimizer is smooth and bounded away from zero and infinity. By the envelope theorem,

$$\frac{\partial L}{\partial a} = -(y-1), \qquad \frac{\partial L}{\partial b} = -(y^{-1} - 1),$$

so the derivatives are uniformly bounded. The Lipschitz bound follows by integrating the pointwise mean-value bound along $\Gamma$. Since $a_\lambda$ and $b_\lambda$ are linear sums of the event intensities, any first-stage rate that controls those rates along the path transfers directly to the action. □

Propositions 15 and 16 are deliberately limited. They cover smooth projections, smoothed loss gaps, and actions for prespecified paths. They do not establish inference for optimized minimum-action paths, hitting probabilities, or rare breakdown probabilities when few breakdowns are observed. Those objects require either smoothing, subsampling, continuity and uniqueness of the minimizing path, or problem-specific large-deviation arguments. The bootstrap tables and action diagnostics should therefore be read as feasible uncertainty and geometry diagnostics, not as substitutes for empirical tail validation.

The same logic also clarifies what aggregate models omit. A vector-autoregressive or conditional-volatility model estimates a projection of $A_X(\lambda_0)$ or $B_X(\lambda_0)$ on a low-dimensional information set. It need not identify the event intensities, the conditional characteristic function, or the higher cumulants generated by those intensities. A Hawkes model may fit event clustering well while leaving the state-dependent demand and supply response poorly measured. The generator approach is useful

when those distinctions matter for distributional fit, economic interpretation, counterfactuals, or rare-event prediction.

**Proposition 17** (Aggregate projections of an event generator)**.** *Let an aggregate variable be* $Y_t = \chi(X_t)$*, and let event* $r$ *change it by the observable increment*

$$\rho_r^Y(X_{t-}) = \chi(X_{t-} + \nu_r) - \chi(X_{t-}).$$

*Under the counting-process target in* (50)*, the first two conditional cumulant rates of* $Y_t$ *are*

$$A_Y(x) = \sum_{r \in \mathcal{R}} \rho_r^Y(x) \lambda_{r,0}(x), \tag{60}$$

$$B_Y(x) = \sum_{r \in \mathcal{R}} \rho_r^Y(x) \rho_r^Y(x)^\top \lambda_{r,0}(x). \tag{61}$$

*Linear projections of* $A_Y(x)$ *generate autoregressive, vector-autoregressive, or vector error-correction mean equations. Projections of* $B_Y(x)$ *generate conditional-variance equations. If* $\lambda_r$ *includes decayed past-event counts, the corresponding event equation is a Hawkes-type intensity.*

*Proof.* Over a short interval, event $r$ occurs with conditional probability $\lambda_{r,0}(X_{t-})\mathrm{d}t + o(\mathrm{d}t)$. Summing the event increments gives the conditional mean rate. Summing outer products of event increments gives the conditional covariance rate. The autoregressive, vector-autoregressive, vector error-correction, and conditional-volatility statements are projections of these conditional moment functions onto lower-dimensional lagged information sets. A decayed past-event count in the intensity is exactly the sufficient statistic used by linear Hawkes specifications. □

Proposition 17 is the local moment version of the broader observable-distribution claim. Once the generator and measurement law imply $F_{Y,\theta}(\cdot \mid x)$, observable forecast targets are evaluated through the scoring rule that elicits the relevant feature: squared loss for means, pinball loss for quantiles, interval scores for prediction intervals, and probability losses for tail events. Proposition 4 states these score-projection targets explicitly. This distinction matters for interpretation: autoregressive, vector-autoregressive, and volatility equations use low-order generator functionals, whereas quantile regression and probability models are direct scoring-rule benchmarks for the observable distribution.

**Proposition 18** (Error transfer to generator projections)**.** *Let* $G(\lambda)$ *be a projection functional*

*such as $A_X(\lambda)$, $B_X(\lambda)$, or a finite-dimensional rare-event score. Suppose $G$ is Lipschitz from the generator norm in* (56) *to the norm used for the projection, and suppose the first-stage event-generator estimator satisfies*

$$\|\hat{\lambda}_K - \lambda_0\|_T = O_p(r_T).$$

*Then*

$$\|G(\hat{\lambda}_K) - G(\lambda_0)\| = O_p(r_T).$$

*If the high-level rate in* (55) *holds, this becomes*

$$\|G(\hat{\lambda}_K) - G(\lambda_0)\| = O_p\left(\sqrt{\frac{K}{T_{\text{eff}}}}\right) + O(a_K)$$

*Alternatively, Assumptions 1 and 9, via Proposition 13, imply*

$$\|G(\hat{\lambda}_K) - G(\lambda_0)\| = O_p\left(\sqrt{\frac{K \log T_{\text{eff}}}{T_{\text{eff}}}}\right) + O(a_K).$$

*Proof.* The first display follows directly from the Lipschitz bound for $G$. The two rate statements substitute the high-level rate in (55) and the primitive rate in Proposition 13. The result is therefore a transfer statement: the projection cannot converge faster than the first-stage generator object unless the projection removes the relevant error directions. □

The same plug-in logic applies to transition-law estimators and to the full observable distribution in (28). For transition data, replace the intensity norm with a metric on transition kernels; any Lipschitz forecast functional inherits the corresponding first-stage rate. Let $K_{\theta,h}$ be the state transition kernel implied by the generator over horizon $h$, let $M_h$ be the measurement or projection law mapping state transitions into the observable, and let $\mathcal{X}_0$ denote the visited state support on which observable forecasts are evaluated. The induced observable distribution is

$$F_{Y,\theta}(y \mid x) = \int M_h((-\infty, y] \mid x, x') K_{\theta,h}(\mathrm{d}x' \mid x).$$

This is a functional of the generator, not a separate reduced-form object.

**Proposition 19** (Plug-in observable distribution)**.** *Suppose $\hat{\theta} \to_p \theta_0$, the transition kernels $K_{\theta,h}(\cdot \mid x)$ are weakly continuous in $\theta$ uniformly on the visited state support, and the measurement kernel*

*$M_h((-\infty, y] \mid x, x')$ is bounded and continuous in $x'$ at each continuity point $y$ of $F_{Y,\theta_0}(\cdot \mid x)$. Then*

$$\sup_{x \in \mathcal{X}_0} |F_{Y,\hat{\theta}}(y \mid x) - F_{Y,\theta_0}(y \mid x)| \to_p 0$$

*at such continuity points. If the conditional density of $Y_{t+h} - Y_t$ is bounded away from zero near a quantile $q_\tau(x)$, then the corresponding generator-implied quantile is consistent. If a scoring rule is continuous in the forecast distribution and dominated by an integrable envelope, the plug-in score converges to the score of the true generator-implied observable distribution.*

*Proof.* Weak continuity of $K_{\theta,h}$ and bounded continuity of the measurement functional imply convergence of the integral defining $F_{Y,\theta}$ by the continuous mapping theorem. Uniformity over $\mathcal{X}_0$ follows from the stated uniform continuity. Quantile consistency is the standard inversion result for distribution functions with positive density at the target quantile. Score convergence follows from dominated convergence applied to the scoring rule evaluated at the plug-in forecast distribution. □

In practice the forecast distribution is often computed by forward simulation, as in (29). This adds a second approximation error, separate from generator estimation error. Conditional on the fitted generator and the held-out state, the simulated observable draws are draws from the plug-in forecast law.

**Proposition 20** (Simulation approximation of the observable forecast law)**.** *Fix the fitted generator, the observable mapping, and a held-out observation $i$. Suppose the simulated observable draws $\tilde{y}_i^{(1)}, \ldots, \tilde{y}_i^{(S_{\mathrm{sim}})}$ are conditionally independent draws from $F_{Y,\hat{\theta}}(\cdot \mid x_i)$. Then, at every continuity point $y$ of $F_{Y,\hat{\theta}}(\cdot \mid x_i)$,*

$$\widehat{F}_{G,i}(y) - F_{Y,\hat{\theta}}(y \mid x_i) = O_p(S_{\mathrm{sim}}^{-1/2})$$

*conditional on the fitted model. For a finite test block, the uniform simulation error over observations is*

$$\max_{i \in \mathcal{T}} |\widehat{F}_{G,i}(y) - F_{Y,\hat{\theta}}(y \mid x_i)| = O_p\left(\sqrt{\frac{\log n_{\mathcal{T}}}{S_{\mathrm{sim}}}}\right)$$

*at the same continuity point. Therefore the simulated forecast distribution is a consistent approximation to the plug-in generator distribution whenever $S_{\mathrm{sim}} \to \infty$, and its error is negligible for test-block averages if $S_{\mathrm{sim}}$ grows faster than $\log n_{\mathcal{T}}$.*

*Proof.* Conditional on the fitted generator and state $x_i$, the simulated distribution function is an

empirical distribution function for independent draws from $F_{Y,\hat{\theta}}(\cdot \mid x_i)$. The pointwise rate follows from the ordinary central limit rate for a Bernoulli average. The finite-block bound follows by applying the same concentration inequality to each held-out observation and taking a union bound. The negligibility statement follows directly from the displayed rate. □

In many applications the measurement law is estimated rather than known. This includes outcome equations with generator signals, residual distributions added to simulated generator paths, and calibrated state-to-outcome mappings. The same plug-in argument applies, but the empirical claim is then joint: forecast scores test the generator and the estimated observable mapping together.

**Proposition 21** (Joint plug-in with estimated observable mapping)**.** *Let $\hat{M}_h$ be an observable measurement law estimated on the training sample, and define*

$$\hat{F}_Y(y \mid x) = \int \hat{M}_h((-\infty, y] \mid x, x') K_{\hat{\theta},h}(\mathrm{d}x' \mid x).$$

*Suppose the conditions of Proposition 19 hold for $K_{\hat{\theta},h}$, and suppose that, for each continuity point $y$,*

$$\sup_{x,x'} \left| \hat{M}_h((-\infty, y] \mid x, x') - M_{0,h}((-\infty, y] \mid x, x') \right| \to_p 0$$

*on the visited state support. Then*

$$\sup_{x \in \mathcal{X}_0} |\hat{F}_Y(y \mid x) - F_{Y,0}(y \mid x)| \to_p 0.$$

*The corresponding quantiles and continuous dominated scores are consistent under the same continuity and positive-density conditions stated in Proposition 19.*

*Proof.* Add and subtract the integral formed with $M_{0,h}$ and $K_{\hat{\theta},h}$. The first difference is bounded by the uniform error of $\hat{M}_h$. The second difference converges by Proposition 19. Quantile and score convergence then follow by the same inversion and dominated-convergence arguments. □

The preceding results can be read as one inferential chain. The ECF step estimates a validation-selected adjustment law. The transition kernel generated by that law is then mapped into an observable forecast distribution. Quantiles, intervals, tail probabilities, and scores are functionals of that distribution. The next proposition states this chain explicitly.

**Proposition 22** (Generator inversion as an observable distribution estimator)**.** *Let $\hat{\alpha} = (\hat{d}, \hat{\rho})$ be selected from a finite library of ECF designs and regularization values as in Proposition 11, and let $\hat{\theta}_{\hat{\alpha}}$ be the corresponding final generator estimate. Suppose the selected population target is $\theta^*$, either because the validation criterion has a separated minimizer or because the selected-rule target is taken as the estimand. Suppose the fixed-design ECF regularity conditions hold at that target and the observable mapping satisfies the continuity conditions in Proposition 19, or the estimated-mapping conditions in Proposition 21. Then, for every continuity point $y$ of the target observable law,*

$$\sup_{x \in \mathcal{X}_0} |F_{Y,\hat{\theta}_{\hat{\alpha}}}(y \mid x) - F_{Y,\theta^*}(y \mid x)| \to_p 0.$$

*If the forecast law is computed by simulation and $S_{\mathrm{sim}}/\log n_{\mathcal{T}} \to \infty$, the simulated forecast distributions are uniformly equivalent to the plug-in forecast laws on the final test block at such continuity points. Consequently, generator-implied means, quantiles, prediction intervals, tail probabilities, and continuous dominated distribution scores converge to the corresponding functionals of the target observable law.*

*If the characteristic design separates the true conditional law and the model is correctly specified, $F_{Y,\theta^*}$ is the true conditional observable distribution for the stated state, horizon, and mapping. If the design is sparse, regularized, or misspecified, $F_{Y,\theta^*}$ is the validation-selected generator projection. In both cases, the final-test score gap estimates the held-out comparison between this generator-implied forecast law and the reduced-form benchmark fixed by the prespecified target, horizon, information rule, selection rule, and score.*

*Proof.* Proposition 11 reduces the adaptive ECF step to a selected finite-grid target, or to the selected-rule target when selection is not separated. Proposition 10 gives convergence of the corresponding generator estimate to that target. Proposition 19, with Proposition 21 when the measurement law is estimated, transfers generator convergence to the observable forecast distribution. Proposition 20 makes simulation error negligible under the stated growth condition. The convergence of distributional functionals and dominated scores follows by the continuous mapping, quantile-inversion, and dominated-convergence arguments already used above. The interpretation of the target follows from Proposition 9 and from the definition of the held-out score gap; Proposition 24 below gives its sampling law. □

The inference layers in Table 45 should be read together with the preceding proposition. Under correct specification and a separating characteristic design, the generator-implied observable law is the true conditional forecast distribution. Under sparse designs, regularization, or misspecification, it is the validation-selected projection that is then compared with reduced-form benchmarks. This separation matters because a tight first-stage generator fit does not by itself imply a precise observable forecast comparison, and a positive score gap is only a claim about the target, horizon, information rule, selection rule, and score fixed before the final test block.

Reduced-form benchmarks must follow the same discipline as the generator. A vector-autoregression lag length, a Gaussian transition specification, a Hawkes projection, a quantile-regression information set, or a rare-event probability model can be chosen on training and validation data. It cannot be chosen after inspecting final-test score gaps. Otherwise the score gap would mix forecast evaluation with benchmark search.

**Proposition 23** (Validation selection of reduced-form benchmarks)**.** *Let $\mathcal{B}_{\text{bench}}$ be a finite library of reduced-form benchmark specifications for a fixed observable target, horizon, and score. For each $b \in \mathcal{B}_{\text{bench}}$, let $\hat{F}_{b,i}$ be the corresponding forecast distribution estimated on the training block, and let $\hat{U}(b)$ be its validation score. If*

$$\sup_{b \in \mathcal{B}_{\text{bench}}} |\hat{U}(b) - U(b)| \to_p 0,$$

*then the validation-selected benchmark $\hat{b} \in \arg\min_{b \in \mathcal{B}_{\text{bench}}} \hat{U}(b)$ satisfies*

$$U(\hat{b}) \leq \min_{b \in \mathcal{B}_{\text{bench}}} U(b) + o_p(1).$$

*If the population validation criterion has a separated minimizer, $\hat{b}$ selects that benchmark with probability approaching one. If several benchmarks are tied, the benchmark in the final score gap is the forecast law generated by the prespecified selection rule.*

*Proof.* The proof is the same finite-library argument as Proposition 11. Uniform convergence over $\mathcal{B}_{\text{bench}}$ makes the validation score of the selected benchmark asymptotically no worse than the best population validation score in the library. Separation gives consistent selection; ties leave selection as part of the benchmark rule. □

The final comparison is a held-out score-gap estimand. Let $\hat{F}_{G,i}$ and $\hat{F}_{B,i}$ be the generator-implied

and benchmark forecast distributions selected before the final test block, with the benchmark selected as in Proposition 23 when a library is used. Let $y_i$ be the realized observable forecast target, and let $S$ be a negatively oriented scoring rule. The reported sample gap is

$$\hat{\Delta}_{S,\mathcal{T}}(G,B) = \frac{1}{n_{\mathcal{T}}} \sum_{i \in \mathcal{T}} \left[ S(\hat{F}_{B,i}, y_i) - S(\hat{F}_{G,i}, y_i) \right].$$

Positive values favor the generator-implied observable distribution.

**Assumption 10** (Held-out score-gap regularity)**.** *The generator, observable mapping, benchmark class, benchmark information rule, benchmark selection rule, penalty choice, and tuning parameters are fixed before the final test block is evaluated. On the test block, the forecast distributions converge uniformly to limiting forecast laws $F^*_{G,i}$ and $F^*_{B,i}$, or their first-stage estimation error is asymptotically negligible for the average score gap. The limiting loss differential*

$$d_i^* = S(F^*_{B,i}, y_i) - S(F^*_{G,i}, y_i)$$

*is stationary and absolutely regular on the test sequence, has a finite $2+\delta$ moment for some $\delta > 0$, and satisfies a block central limit theorem with long-run variance $\Lambda_S$.*

**Proposition 24** (Held-out score-gap inference)**.** *Under Assumption 10,*

$$\hat{\Delta}_{S,\mathcal{T}}(G,B) \to_p \Delta_S(G,B) = \mathbb{E}[d_i^*],$$

*and*

$$\sqrt{n_{\mathcal{T}}}\{\hat{\Delta}_{S,\mathcal{T}}(G,B) - \Delta_S(G,B)\} \Rightarrow N(0, \Lambda_S).$$

*If a moving-block bootstrap is applied to the held-out loss-differential sequence with block length $L \to \infty$ and $L/n_{\mathcal{T}} \to 0$, and the bootstrap consistently estimates the block central limit law for $d_i^*$, then the bootstrap distribution consistently estimates the sampling law of the held-out score gap.*

*Proof.* Uniform convergence of the forecast distributions, or asymptotic negligibility of their estimation error, reduces the sample gap to the average of $d_i^*$ up to $o_p(n_{\mathcal{T}}^{-1/2})$. The law of large numbers for absolutely regular sequences gives consistency, and the stated block central limit theorem gives the normal limit. The moving-block bootstrap conclusion follows from the assumed bootstrap consistency for the same dependent loss-differential sequence. □

The score gap in (30) is therefore an evaluation estimand, not an additional identifying assumption. Once the generator, the observable mapping, and the benchmark information and selection rules are fixed before the final test block, the held-out average score gap estimates which forecast distribution is better for that target and horizon. Proposition 24 gives the corresponding uncertainty calculation. This is the object an empirical application should report alongside first-stage generator diagnostics.

## 10 Implementation and Evaluation Protocols

The implementation protocol is domain-agnostic. It starts from event or transition data, defines the state observed before the adjustment, estimates the conditional adjustment law identified by the observation regime, and evaluates the observable conditional distributions induced by that law. With event histories this object can be a local generator on the exposed support. With transition panels it is first a conditional transition law at the observed horizons; it becomes a generator claim only under the semigroup, multi-horizon, or injectivity conditions stated in Section 9. The protocol does not choose the empirical domain. It states what any domain must supply: adjustment events or transitions, state variables, ECF frequencies, instruments, split rule, regularization grid, observable forecast targets, and reduced-form benchmarks fixed before final testing.

Two input regimes make the protocol concrete without making either one the selected domain. The first is a lower-frequency transition path. In a dairy processor panel, for example, the existing German milk-market data contain monthly processor payout prices, processor production direction, throughput, weekly butter and skimmed-milk-powder cash prices, and EEX futures signals. Such data are enough to study the conditional transition law of monthly payout adjustments,

$$\Delta p_{it} \mid x_{it},$$

where $x_{it}$ contains lagged disequilibrium, processor type, product-market exposure, throughput, and benchmark-price changes. The observable target is the payout-price adjustment itself, possibly augmented by delivered quantities or product-output shares when those variables are available. This is a transition-data version of the same distributional forecasting problem. It can support generator language only when the transition model is linked across horizons or otherwise made injective; without that extra structure, the validated claim is about the conditional transition law and the observable distributions it implies.

Table 48 states the same point as a feasibility template. A processor panel is useful for the framework when it has a well-defined unit, horizon, outcome, state vector, and benchmark set. It is not an event-history application. Used conservatively, it can test whether adjustment-state information improves the forecast distribution of monthly payout-price changes relative to error-correction, autoregressive, Gaussian, and quantile-regression benchmarks. It should not be described as identifying a continuous-time adjustment generator unless additional horizon-linking or injectivity restrictions are imposed.

Table 48: Transition-panel route for monthly processor-price adjustments.

The table summarizes how a lower-frequency processor panel can enter generator inversion as a transition-law application. It separates feasible forecast objects from the stronger generator claim that would require additional identifying structure.

| Design object | Feasible construction | Claim boundary |
|---|---|---|
| *Observed transition panel* | | |
| Unit and horizon | Monthly processor records for thirteen German dairy processors, available under monthly-mean and end-of-month aggregation rules. | The directly observed object is a one-month transition panel, not an event history. |
| Observable outcome | Processor comparison-price adjustment, measured in cents per kilogram of raw milk on a standardized fat-and-protein basis. | The forecast target is the conditional distribution of payout-price changes, not a high-frequency order-flow object. |
| State variables | Lagged processor price, production direction, processor size, spot milk price, butter and skimmed-milk-powder cash prices, and butter and skimmed-milk-powder futures signals. | State design must be fixed before validation and final testing; aggregation rule is part of the design. |
| *Generator-inversion interpretation* | | |
| Adjustment law | Conditional transition law for monthly processor-price adjustments given processor type, size, lagged disequilibrium, and benchmark-market signals. | With one observed horizon, the validated object is a transition law. A generator claim requires event histories, linked horizons, or an injective transition-to-generator parameterization. |
| Observable distribution | Forecast distribution for monthly payout-price changes, constructed from the fitted transition law and evaluated on a final held-out block. | The paper can score means, quantiles, intervals, tails, and full-distribution approximations without claiming continuous-time generator identification. |
| Benchmarks | Error-correction model, vector autoregression, Gaussian transition model, direct quantile regression, and rare-adjustment probability model. | All benchmark information sets, lag choices, quantile grids, and selection rules must be fixed before final-test scoring. |

The second input regime is an event-history path. A high-frequency order-message stream is useful as a test case because arrivals and removals are observed directly, so the event-to-generator mapping can be audited. The minimum raw record is an ordered message stream with a timestamp, side, action, price state, and preferably the outstanding book state before the event. In the order-message schema, each message is mapped into one of four generator events:

$$\{D^+, D^-, S^+, S^-\} = \{\text{demand arrival}, \text{demand cancellation}, \text{supply arrival}, \text{supply cancellation}\}.$$

The pre-event state then contains the demand depth, supply depth, relative price, imbalance, total depth, and any additional covariates used in the chosen basis. This mapping is deliberately simple. Its purpose is to make one event-history input layer auditable. A different application would replace these events and state variables with the economically relevant adjustments, such as price revisions, production reallocations, inventory changes, defaults, job transitions, or investment adjustments. The estimator and final scoring rule do not depend on the order-book interpretation.

The optional event-history construction has three steps. First, the input layer records candidate message files by date and feed and leaves large raw downloads outside ordinary builds. Second, the normalization layer reads decoded order-message records and converts them into the canonical event schema. Order-by-order displayed-depth messages identify additions, executions, cancellations, and deletions. Additions are arrivals; executions, cancellations, and deletions are removals. If only price-level displayed-depth updates are available, positive displayed-depth changes are interpreted as arrivals and negative changes as removals. This fallback is a net reconstruction at the price level, not an order-by-order history. Third, the validation step checks the canonical schema and writes the event history used by the estimator. A lower-frequency transition application requires the analogous mapping from panel records to canonical adjustment increments.

The benchmark step is common across input regimes. It first defines the state, the observable increment to be explained, the frequency grid, the distribution-forecast quantiles, and the state instruments before looking at the final test window. In an event-history path, the initialization and comparison model is the point-process likelihood for event intensities. In a transition-panel path, the comparison model can be an error-correction model, vector autoregression, direct quantile regression, or parametric transition density for the observed increment. The conditional ECF estimator is then selected on the validation window with a ridge grid around the likelihood or transition anchor and evaluated on the untouched test split. Final reporting uses ECF losses, likelihood or transition-fit benchmarks where available, conditional moments, observable adjustment, volatility, rare-state prediction, quantile and interval scores for observable distribution forecasts, and the value of adding generator-implied adjustment signals to ordinary aggregate forecasts. If no valid event history or transition panel is available, the paper reports protocol readiness rather than a benchmark result.

For transition panels, the final report should be explicit about the strength of the claim. If the panel observes one fixed horizon, the directly validated object is the conditional transition law for that horizon. The paper can still evaluate its observable distribution forecasts, because those

forecasts are functionals of the same transition law. A stronger generator claim requires either event histories, multiple horizons connected by a common semigroup, or a parameterization that makes the transition-to-generator map injective on the visited support. This reporting rule prevents the lower-frequency application from borrowing more identification than the data provide.

The executable event-history rule is fixed before seeing the final test split. Table 49 reports the main protocol items for that pilot design. The important design choice is the separation between basis selection and final evaluation. Candidate bases are estimated on the first 60 percent of the event stream and selected on the next 20 percent. The selected state-based generator and the aggregate benchmarks are then evaluated on the final 20 percent. Distribution forecasts use a prespecified quantile grid and a fixed number of generator simulation paths, so the direct comparison with quantile regression is not tuned on the test sample. The transition-panel version follows the same logic with calendar or unit blocks: choose the state basis, ECF frequency grid, and distributional scoring design on the training and validation periods, then report ECF loss and benchmark gaps only on the final period. This does not solve the empirical-data problem, but it prevents the benchmark table from choosing a basis on the same observations used to report the paper's central losses.

Table 50 records the corresponding lower-frequency transition design. Its purpose is to make transition-panel applications a first-class route through the same framework. When events are latent, the estimator still targets the conditional law of the increment. The table therefore separates the transition law, the measurement law that turns state transitions into observable forecasts, and the claim boundary. With one observed horizon the directly validated object is a transition-law forecast; a generator interpretation requires event histories, semigroup-linked horizons, or injectivity. The anchor changes from an event likelihood to a transition fit, but the ECF moments, validation logic, observable distribution forecasts, and benchmark scores remain the same.

When the local processor panel is available, the transition route can be run as a bounded diagnostic. The diagnostic treats processor-month records as one-month transitions and compares a regularized characteristic estimator with a Gaussian transition anchor and a direct quantile benchmark. It is not used to claim event-level generator identification. Its role is to show whether the same estimator-score chain can be executed on a lower-frequency panel and whether the resulting observable forecast evidence agrees with the characteristic-function evidence.

The reported columns follow the definitions in Sections 4 and 5. The sample split lists the train-

Table 49: Prespecified protocol for executable event-history benchmarks.

The table lists the train-validation-test split, model-selection rule, event-count threshold, ECF frequency grids, instruments, regularization grid, and weighting choices used for an executable event-history benchmark design.

| Protocol item | Value | Role |
|---|---|---|
| Training share | 0.60 | Estimate candidate event-intensity models. |
| Validation share | 0.20 | Select the event-history basis by prespecified validation criterion. |
| Test share | 0.20 | Hold out final benchmark evaluation from basis selection. |
| Minimum event count | 200 | Minimum normalized events before event-history benchmarks are reported. |
| Candidate event bases | Linear state, Quadratic state, Spline state, Hawkes-style | Candidate bases for event-likelihood validation. |
| Candidate state bases | Linear state, Quadratic state, Spline state | Candidate bases for aggregate generator projections. |
| Selection criterion | Mean validation event log likelihood | Primary validation criterion for basis choice. |
| Rare-event horizon | 20 | Forward event horizon for rare liquidity-state labels. |
| Low-depth cutoff | 10th percentile of training-sample total depth | Training-sample depth quantile defining low-depth states. |
| Extreme-imbalance cutoff | 90th percentile of training-sample absolute imbalance | Training-sample absolute-imbalance quantile defining extreme imbalance. |
| Final held-out reports | Event likelihood; conditional characteristic-function loss; conditional-moment loss; price-change loss; volatility loss; rare-event probability and ranking; time-series value added by generator signals; observable distribution forecasts; quantile-regression benchmark | Final held-out reporting targets. |
| Low-frequency ECF grid | 0.25, 0.50, 0.75 | Low-frequency ECF grid emphasizing drift and variance restrictions. |
| Medium-frequency ECF grid | 1.00, 1.50, 2.00 | Medium-frequency ECF grid emphasizing discreteness and higher-cumulant restrictions. |
| ECF instruments | Excess demand, Pre-event price, Pre-event imbalance, Log pre-event depth | Standardized state instruments; an intercept is added automatically. |
| ECF block length | 25 | Block length for long-run covariance weighting and ECF loss-gap bootstrap. |
| ECF weighting | Identity initialization; block long-run covariance in step two | Use identity weighting for initialization and block long-run covariance for the second ECF step. |
| ECF regularization | Anchor fit, 0, 0.00001, 0.0001, 0.001, 0.01, 0.1 | Validation-selected ridge penalties around the likelihood or transition anchor; anchor keeps the fitted generator when ECF moments do not improve out of sample. |
| ECF bootstrap replications | 99 | Moving-block bootstrap replications for held-out ECF loss-gap intervals. |
| ECF pilot row cap | 800 | Bounded pilot-row cap for ECF optimization; increase for scale runs after profiling. |
| Distribution forecast quantiles | 0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95 | Quantiles used to evaluate observable distribution forecasts and quantile-regression benchmarks. |
| Generator forecast paths | 250 | Monte Carlo paths per held-out state for generator-implied observable distribution forecasts. |

Table 50: Protocol for lower-frequency transition-panel applications.

The table states how the same generator-inversion design is implemented when individual adjustment events are latent and the data record transitions between observed states. Panel headers separate data construction, estimator design, and observable evaluation.

| Protocol item | Rule | Role |
|---|---|---|
| *Data construction* | | |
| Observation unit | Economic unit by period | Defines the transition record, such as processor-month, firm-quarter, or borrower-month. |
| Initial state | Lagged outcome, disequilibrium, exposure, type, and aggregate controls | Conditions the transition law on information available before the adjustment. |
| Adjustment increment | Future state minus initial state over the chosen horizon | Supplies the realized increment used in characteristic moments and forecast scoring. |
| Observed horizon | One or more prespecified forecast horizons | Determines whether the directly validated object is a single-horizon transition law or a horizon-linked semigroup restriction. |
| Observable target | Price, quantity, inventory, state, duration, or stress outcome | Defines the researcher-facing outcome distribution implied by the generator. |
| Measurement law | Deterministic state-to-outcome map or estimated projection law | Translates transition states into the observable forecast distribution that is scored. |
| *Estimator design* | | |
| Anchor model | Transition likelihood, error-correction model, vector autoregression, or prespecified transition fit | Provides initialization, comparison, and regularization center when individual events are latent. |
| Frequency grid | Low and medium frequencies chosen before final testing | Low frequencies emphasize mean and variance; medium frequencies test discreteness, skewness, and tail shape. |
| State instruments | Standardized lags, disequilibrium measures, exposure variables, and aggregate controls | Convert unconditional characteristic moments into conditional restrictions. |
| Split rule | Train, validation, and final test blocks by calendar time or unit clusters | Separates estimation, ECF regularization choice, and final reporting. |
| Regularization grid | Anchor plus finite ridge grid around the anchor | Allows the ECF step to move only when validation characteristic loss supports it. |
| Generator-identification condition | Event histories, multiple horizons with semigroup restrictions, or an injective transition-to-generator parameterization | States when transition panels identify a generator rather than only the observed transition law. |
| Claim boundary | Transition-law forecast unless event histories, semigroup-linked horizons, or injectivity justify a generator interpretation | Prevents lower-frequency panels from supporting stronger path or counterfactual claims than the data identify. |
| *Observable evaluation* | | |
| Distribution forecast | Simulated or analytic predictive distribution for each held-out observation | Produces quantiles, intervals, tail probabilities, and point projections from one conditional law. |
| Reduced-form benchmarks | Vector autoregression, Gaussian transition model, quantile regression, and rare-event probability model | Tests whether generator information improves on direct observable forecasts. |
| Reported scores | Characteristic loss, pinball loss, interval score, tail probability loss, and rare-event scores | Evaluates the conditional distribution rather than only the conditional mean. |

ing, validation, and final-test observations in that order. Conditional characteristic loss is $10^4$ times (18); unconditional characteristic loss repeats the same calculation with only the intercept instrument. The weighted characteristic criterion is the unscaled quadratic moment criterion $g'Wg$ for the reported model and frequency band. The transition diagnostic uses the quantile grid $\{0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95\}$. Its lower-tail threshold is the tenth percentile of training-and-validation price changes, fixed before the final test block is scored. The distribution table reports the average pinball loss, the equal-weight quantile-grid approximation to the continuous ranked probability score, observable characteristic-function loss from (31), the lower-tail probability loss, and the share of crossed adjacent quantiles. The observable characteristic loss is multiplied by $10^4$; for the direct quantile benchmark it uses the displayed quantile grid as a finite forecast-distribution approximation. The gap table reports benchmark loss minus regularized-transition-estimator loss for the same criteria, so positive entries favor the generator-implied observable distribution.

The transition-panel tables use the same audit rule as the simulation tables. The caption names the target, the description states the interpretation of the entries, and Appendix C links recurring measures to their formula, block, scale, and sign convention. This is especially important for the processor diagnostic because characteristic losses and observable forecast scores answer different questions: the first checks the fitted transition law, while the second checks the induced distribution of payout-price changes against direct reduced-form benchmarks.

Table 51: Monthly processor-price transition sample.

The table reports the processor count, sample window, transition count, train-validation-test split, and selected characteristic design for the monthly processor-price transition diagnostic.

| Aggregation rule | Processors | Months | Transitions | Split | Selected design |
|---|---|---|---|---|---|
| End-of-month aggregation | 13 | 2014-01-01–2019-10-01 | 845 | 507/169/169 | Low-frequency transition |
| Monthly-mean aggregation | 13 | 2014-01-01–2019-11-01 | 923 | 546/182/195 | Low-frequency transition |

Table 52: Validation-selected characteristic designs for processor-price transitions.

The table reports the validation loss, moment count, selected penalty, and final decision for candidate characteristic designs. Moment count is the number of stacked real and imaginary characteristic moments after combining frequencies and standardized state instruments. Panel headers separate the two temporal aggregation rules.

| Candidate design | Validation loss | Moment count | Penalty | Decision |
|---|---|---|---|---|
| *End-of-month aggregation* | | | | |
| Low-and-medium transition | 1.4139 | 96 | 0.01 | |
| Low-frequency transition | 0.1493 | 36 | 0.01 | Selected |
| *Monthly-mean aggregation* | | | | |
| Low-and-medium transition | 14.1531 | 96 | Anchor | |
| Low-frequency transition | 0.2442 | 36 | 0.01 | Selected |

Table 53: Processor-price transition characteristic losses.

The table compares held-out conditional and unconditional characteristic-function losses for the regularized transition estimator, its Gaussian anchor, and an unconditional Gaussian transition benchmark. Conditional losses use training-standardized state instruments, while unconditional losses use only the intercept. Losses are scaled by $10^4$; lower values indicate better held-out characteristic fit.

| Model | Conditional characteristic loss $\times 10^4$ | Unconditional characteristic loss $\times 10^4$ | Weighted characteristic criterion |
|---|---|---|---|
| *End-of-month aggregation* | | | |
| *Low frequency band* | | | |
| Gaussian transition anchor | 110.268 | 267.873 | 0.1985 |
| Regularized transition characteristic estimator | 87.733 | 211.507 | 0.1579 |
| Unconditional Gaussian transition | 106.645 | 344.903 | 0.1920 |
| *Medium frequency band* | | | |
| Gaussian transition anchor | 584.511 | 1804.592 | 1.0521 |
| Regularized transition characteristic estimator | 462.313 | 1475.115 | 0.8322 |
| Unconditional Gaussian transition | 907.516 | 3027.228 | 1.6335 |
| *Monthly-mean aggregation* | | | |
| *Low frequency band* | | | |
| Gaussian transition anchor | 124.488 | 329.976 | 0.2241 |
| Regularized transition characteristic estimator | 6.736 | 20.427 | 0.0121 |
| Unconditional Gaussian transition | 162.000 | 414.890 | 0.2916 |
| *Medium frequency band* | | | |
| Gaussian transition anchor | 650.918 | 1826.363 | 1.1717 |
| Regularized transition characteristic estimator | 172.187 | 524.826 | 0.3099 |
| Unconditional Gaussian transition | 1118.516 | 3064.488 | 2.0133 |

Table 54: Processor-price transition distribution scores.

The table evaluates the observable distribution of monthly payout-price changes. It reports final-test average pinball loss, a quantile-grid approximation to the continuous ranked probability score, observable characteristic-function loss, tail-probability loss, and quantile-crossing share for the regularized transition estimator, Gaussian transition anchor, and direct quantile benchmark. The observable characteristic loss is scaled by $10^4$, and the tail-probability threshold is fixed from the training-and-validation block.

| Model | Average pinball loss | Continuous ranked probability score approximation | Observable characteristic-function loss $\times 10^4$ | Tail probability loss | Crossing share |
|---|---|---|---|---|---|
| *End-of-month aggregation* | | | | | |
| Direct quantile benchmark | 0.1325 | 0.2650 | 341.4513 | 0.0152 | 1.2% |
| Gaussian transition anchor | 0.1415 | 0.2830 | 255.4858 | 0.0131 | 0.0% |
| Regularized transition characteristic estimator | 0.1811 | 0.3623 | 200.2336 | 0.0136 | 0.0% |
| *Monthly-mean aggregation* | | | | | |
| Direct quantile benchmark | 0.1521 | 0.3041 | 476.7486 | 0.0226 | 5.6% |
| Gaussian transition anchor | 0.1504 | 0.3007 | 251.1817 | 0.0195 | 0.0% |
| Regularized transition characteristic estimator | 0.1340 | 0.2681 | 26.3335 | 0.0231 | 0.0% |

Table 55: Processor-price transition distribution score gaps.

The table reports final-test score gaps for monthly payout-price changes. Each panel fixes an aggregation rule and benchmark; entries equal benchmark loss minus regularized transition-estimator loss, so positive gaps favor the generator-implied observable distribution. Observable characteristic-function loss entries are scaled by $10^4$, matching the scale in the distribution-score table.

| Criterion | Generator gap | Generator loss | Benchmark loss |
|---|---|---|---|
| *End-of-month aggregation: Direct quantile benchmark* | | | |
| Average pinball loss | -0.0486 | 0.1811 | 0.1325 |
| Continuous ranked probability score approximation | -0.0972 | 0.3623 | 0.2650 |
| Observable characteristic-function loss $\times 10^4$ | 141.2177 | 200.2336 | 341.4513 |
| Tail probability loss | 0.0016 | 0.0136 | 0.0152 |
| *End-of-month aggregation: Gaussian transition anchor* | | | |
| Average pinball loss | -0.0396 | 0.1811 | 0.1415 |
| Continuous ranked probability score approximation | -0.0793 | 0.3623 | 0.2830 |
| Observable characteristic-function loss $\times 10^4$ | 55.2522 | 200.2336 | 255.4858 |
| Tail probability loss | -0.0005 | 0.0136 | 0.0131 |
| *Monthly-mean aggregation: Direct quantile benchmark* | | | |
| Average pinball loss | 0.0180 | 0.1340 | 0.1521 |
| Continuous ranked probability score approximation | 0.0361 | 0.2681 | 0.3041 |
| Observable characteristic-function loss $\times 10^4$ | 450.4151 | 26.3335 | 476.7486 |
| Tail probability loss | -0.0005 | 0.0231 | 0.0226 |
| *Monthly-mean aggregation: Gaussian transition anchor* | | | |
| Average pinball loss | 0.0163 | 0.1340 | 0.1504 |
| Continuous ranked probability score approximation | 0.0327 | 0.2681 | 0.3007 |
| Observable characteristic-function loss $\times 10^4$ | 224.8482 | 26.3335 | 251.1817 |
| Tail probability loss | -0.0036 | 0.0231 | 0.0195 |

Table 56: Processor-price transition prediction intervals.

The table compares final-test coverage, average width, and interval scores for central prediction intervals of monthly payout-price changes. The 80 percent interval uses the 0.10 and 0.90 quantiles; the 90 percent interval uses the 0.05 and 0.95 quantiles. Panel headers separate aggregation rules; lower interval scores indicate better calibrated and sharper intervals.

| Interval | Model | Coverage | Width | Interval score |
|---|---|---|---|---|
| *End-of-month aggregation* | | | | |
| 80 percent interval | Direct quantile benchmark | 95.9% | 2.1716 | 2.3438 |
| 80 percent interval | Gaussian transition anchor | 91.7% | 2.0967 | 2.3799 |
| 80 percent interval | Regularized transition characteristic estimator | 74.0% | 1.6230 | 2.9549 |
| 90 percent interval | Direct quantile benchmark | 98.2% | 2.9148 | 3.1365 |
| 90 percent interval | Gaussian transition anchor | 96.4% | 2.6911 | 2.9489 |
| 90 percent interval | Regularized transition characteristic estimator | 82.2% | 2.0830 | 3.7509 |
| *Monthly-mean aggregation* | | | | |
| 80 percent interval | Direct quantile benchmark | 95.4% | 2.4802 | 2.6619 |
| 80 percent interval | Gaussian transition anchor | 93.8% | 2.2650 | 2.5422 |
| 80 percent interval | Regularized transition characteristic estimator | 85.6% | 1.6700 | 2.1954 |
| 90 percent interval | Direct quantile benchmark | 96.9% | 3.1316 | 3.3517 |
| 90 percent interval | Gaussian transition anchor | 97.4% | 2.9071 | 3.2046 |
| 90 percent interval | Regularized transition characteristic estimator | 94.4% | 2.1434 | 2.7525 |

The application-readiness layer reads the benchmark audit and writes a separate status table. This prevents missing event or transition data from being interpreted as a zero or failed estimate. Table 57 records whether a normalized event history or transition panel and final benchmark outputs are available. When a benchmark has been run on valid input records, the generated report also includes sample diagnostics, validation fit by basis, final-test likelihood or transition-fit evidence, held-out projection losses, and the value of adding generator signals to aggregate time-series forecasts.

Table 57: Application-readiness status.

The table reports whether the protocol and the normalized event-history or transition-panel input needed for benchmark estimation are available in the current build.

| Protocol item | Status |
|---|---|
| Protocol | Available |
| Normalized event or transition history | found |
| Train, validation, test rows | 9486/3160/3162 |
| Selected state basis | Linear state |
| Final likelihood or transition fit | Available |
| Characteristic ECF losses | Available |
| Held-out rare-event positives | 0 |
| Benchmark status | Ready |

Table 58: Empirical sample and split diagnostics.

The table reports the empirical sample size, event composition, and train-validation-test split after the event history has been normalized.

| Diagnostic | Value | Interpretation |
|---|---|---|
| Normalized events | 15808 | Rows in the normalized event history used by the benchmark. |
| Observed time span | 30445.1210 | Elapsed event-clock time from the first to the last normalized message. |
| Train / validation / test rows | 9486 / 3160 / 3162 | Rows assigned by the prespecified chronological split. |
| Median waiting time | 0.4510 | Median duration between consecutive messages. |
| 99th percentile waiting time | 23.6284 | Upper-tail duration between consecutive messages. |
| Nonpositive waiting times | 0 | Rows with invalid or nonpositive durations after normalization. |
| Demand arrivals | 4551 | Events that add demand depth. |
| Demand cancellations | 3374 | Events that remove demand depth. |
| Supply arrivals | 4436 | Events that add supply depth. |
| Supply cancellations | 3447 | Events that remove supply depth. |
| Test rare-event positives | 0 | Held-out observations labeled as future low-depth or high-imbalance states. |

The protocol mirrors the simulation benchmarks. First, estimate the generator or transition law under prespecified basis families. Second, choose or average bases using rolling-block validation, with the basis-selection rule fixed before final testing. Third, compare generator and process-first alter-

Table 59: Empirical validation likelihoods by event-intensity basis.

The table compares candidate event-intensity bases on the validation sample, using event likelihood and, when available, intensity-recovery diagnostics.

| Basis | Validation log likelihood per interval |
|---|---|
| Linear state | -1750.6740 |
| Quadratic state | -1767.3360 |
| Hawkes-style | -2036.8950 |
| Spline state | NA |

Table 60: Empirical final-test event likelihood by event-intensity basis.

The table reports final held-out event likelihoods for the selected empirical basis and any comparison bases that remain in the benchmark set.

| Basis | Role | Test log likelihood per interval |
|---|---|---|
| Hawkes-style | Hawkes-style event benchmark | -1724.5520 |
| Quadratic state | Other prespecified basis | -1978.2070 |
| Linear state | Selected state generator | -1984.0940 |
| Spline state | Other prespecified basis | NA |

Table 61: Empirical conditional characteristic-function losses, scaled by $10^4$.

The table reports conditional and unconditional characteristic-function losses by frequency band and model, allowing the empirical generator to be compared with aggregate distributional projections.

| Model | Conditional characteristic loss | Unconditional characteristic loss | Weighted characteristic criterion |
|---|---|---|---|
| *Low (0.25–0.75)* | | | |
| Likelihood jump generator | 24.167 | 44.712 | 0.0363 |
| Regularized characteristic-function jump generator | 20.706 | 37.875 | 0.0311 |
| Gaussian truncation | 24.550 | 45.567 | 0.0368 |
| Aggregate Gaussian projection | 1083.971 | 1834.210 | 1.6260 |
| Hawkes-style generator | 89.567 | 161.969 | 0.1344 |
| *Medium (1.00–2.00)* | | | |
| Likelihood jump generator | 172.034 | 260.567 | 0.2581 |
| Regularized characteristic-function jump generator | 173.271 | 262.302 | 0.2599 |
| Gaussian truncation | 169.565 | 256.907 | 0.2543 |
| Aggregate Gaussian projection | 236.122 | 423.264 | 0.3542 |
| Hawkes-style generator | 166.310 | 253.396 | 0.2495 |

Table 62: Empirical ECF loss gaps relative to the anchor generator.

The table reports held-out ECF loss gaps and bootstrap intervals relative to the likelihood or transition anchor, separated by frequency band.

| Model vs likelihood | Generator gap $\times 10^4$ | 95% interval | Positive share |
|---|---|---|---|
| *Low frequency band* | | | |
| Aggregate Gaussian projection | -1059.804 | [-1263.985, -859.821] | 0% |
| Gaussian truncation | -0.383 | [-0.648, 0.017] | 5% |
| Hawkes-style generator | -65.400 | [-86.897, -39.249] | 0% |
| Regularized characteristic-function jump generator | 3.461 | [1.101, 5.581] | 100% |
| *Medium frequency band* | | | |
| Aggregate Gaussian projection | -64.088 | [-171.349, 47.929] | 14% |
| Gaussian truncation | 2.469 | [1.606, 3.539] | 100% |
| Hawkes-style generator | 5.724 | [2.879, 9.470] | 100% |
| Regularized characteristic-function jump generator | -1.237 | [-1.809, -0.793] | 0% |

Table 63: Empirical held-out generator and benchmark losses.

The table compares final held-out losses for generator-based projections and reduced-form benchmark models across the empirical targets.

| Model | Target | Metric | Test val |
|---|---|---|---|
| Generator signals | Conditional moments | Excess-demand drift error | 598.95 |
| Time-series benchmark | Conditional moments | Excess-demand drift error | 548.04 |
| Generator signals | Conditional moments | Second-moment error | 4557115.00 |
| Time-series benchmark | Conditional moments | Second-moment error | 4264282.00 |
| Generator signals | Price adjustment | Price-change error | 10.04 |
| Vector-autoregression benchmark | Price adjustment | Price-change error | 10.00 |
| Generator signals | Volatility | Squared price-change error | 1306.71 |
| Conditional-volatility benchmark | Volatility | Squared price-change error | 1285.76 |
| Generator rare-event model | Rare liquidity state | Area under receiver-operating-characteristic curve | N |
| Time-series rare-event model | Rare liquidity state | Area under receiver-operating-characteristic curve | N |
| Generator rare-event model | Rare liquidity state | Brier score | N |
| Time-series rare-event model | Rare liquidity state | Brier score | N |

Table 64: Empirical tail concentration of rare liquidity-state forecasts.

The table reports whether observations ranked as highest risk by each empirical model contain a higher concentration of future rare liquidity states.

| Model | Tail diagnostic | Value |
|---|---|---|
| Generator rare-event model | Held-out observations | NA |
| Generator rare-event model | Breakdown rate | NA% |
| Generator rare-event model | Highest-risk decile breakdown rate | NA% |
| Generator rare-event model | Highest-risk decile lift | NA |
| Generator rare-event model | Highest-risk decile positives | NA |
| Time-series rare-event model | Held-out observations | NA |
| Time-series rare-event model | Breakdown rate | NA% |
| Time-series rare-event model | Highest-risk decile breakdown rate | NA% |
| Time-series rare-event model | Highest-risk decile lift | NA |
| Time-series rare-event model | Highest-risk decile positives | NA |
| Time series plus generator signals | Held-out observations | NA |
| Time series plus generator signals | Breakdown rate | NA% |
| Time series plus generator signals | Highest-risk decile breakdown rate | NA% |
| Time series plus generator signals | Highest-risk decile lift | NA |
| Time series plus generator signals | Highest-risk decile positives | NA |

Table 65: Empirical value of adding generator signals to time-series forecasts.

The table compares aggregate time-series forecasts with and without generator-implied signals, reporting the held-out gain for each empirical forecast target.

| Forecast target | Loss or score | Time series only | With generator signals | Generator gain |
|---|---|---|---|---|
| Price changes | Root mean squared error | 10.0079 | 10.0059 | 0.0020 |
| Squared price changes | Root mean squared error | 1141.6027 | 1136.7188 | 4.8839 |
| Rare-stress probability | Brier score | NA | NA | NA |
| Rare-stress ranking | Area under receiver-operating-characteristic curve | NA | NA | NA |

natives on held-out criteria: likelihood or transition fit where available, conditional characteristic-function loss, conditional moment fit, volatility forecast loss, rare-state prediction, and observable distribution scores against quantile-regression benchmarks. Fourth, report where the generator improves measurement and where aggregate models are already sufficient. Fifth, report the executable checks that produced the evidence: normalization checks for event or transition records, ECF estimator tests, observable distribution forecast checks, table-generation checks, and a successful paper build. This last step is essential: the evidence should identify the information content of adjustment data, not merely show that a more detailed model can be estimated or that the simulation laboratory is internally consistent.

# 11 Conclusion

This paper develops generator inversion as a distributional econometric framework. The estimand is the conditional adjustment generator: the law that maps current states into future events or transitions. The estimator is an adaptive regularized conditional empirical-characteristic-function minimum-distance procedure that selects its characteristic design and regularization on validation data while retaining likelihood or transition fitting as an anchor. The bridge is a transition kernel together with a measurement or projection law that maps adjustment states into observed outcomes. The statistical output is the conditional distribution of observables implied by that selected generator and measurement law, not the generator as a latent object for its own sake.

The central message is that a dynamic generator is useful only through its observable implications. Evaluated at imaginary momentum, the generator's Hamiltonian is a local conditional characteristic exponent. That object links event likelihoods, characteristic-function restrictions, observable outcome projections, cumulants, Gaussian diffusion approximations, quantiles, prediction intervals, tail probabilities, and rare-event actions. It also clarifies how standard econometric models enter the framework. Depending on the observation regime and information set, vector autoregressions, Gaussian transition models, volatility equations, Hawkes models, and quantile regressions are moment projections, likelihood anchors, score projections, or direct benchmarks for the observable distributions induced by the generator and measurement law.

The synthesis result makes this message econometric rather than only interpretive. Under the stated empirical-characteristic-function regularity, validation, and mapping conditions, the selected

generator and measurement law deliver a consistent observable forecast distribution. Under misspecification, the estimand is explicit: the validation-selected generator projection whose observable distribution is compared with prespecified reduced-form benchmarks on the final test block.

The simulation evidence supports a calibrated methodological claim. It demonstrates how a state-dependent jump generator can be estimated, regularized with characteristic-function restrictions, translated into observable predictive distributions, and compared with direct reduced-form forecasts. In the current calibration, the generator improves some diagnostics and does not dominate others. That is the right standard for the paper: generator inversion should be treated as a disciplined distributional representation, not as a guarantee that richer adjustment data always beat direct observable models.

The contribution is therefore fivefold. First, the paper defines a conditional adjustment generator as an econometric estimand for event histories and transition panels. Second, it estimates this object with regularized conditional characteristic moments, anchored when appropriate by likelihood or transition fitting. Third, it maps the fitted adjustment law into a predictive distribution for observables. Fourth, it evaluates that distribution with adjustment and observable characteristic-function losses, quantile losses, interval scores, tail-probability losses, rare-event ranking scores, and score gaps. Fifth, it interprets standard econometric models as moment projections, likelihood anchors, score projections, or direct benchmarks for the same observable forecast problem.

The paper does not claim that richer event models always win. It provides a framework for measuring where adjustment data contain predictive distributional information and where direct reduced-form models are already sufficient. The value condition is sharp: under proper scoring rules, a generator can add observable forecast value only when the adjustment information changes the conditional distribution and the estimator recovers that change out of sample. Future empirical work can apply the same design to order books, processor panels, credit transitions, labor-market flows, firm dynamics, or any setting with event or transition histories. Future theory should strengthen inference for optimized tail objects such as hitting probabilities and minimum-action paths. The present paper supplies the estimable object and the diagnostic discipline: estimate the conditional adjustment law, translate it into observable conditional distributions, and evaluate those distributions against transparent reduced-form benchmarks.

# A Worked Example: Liquidity Birth–Death System

The minimal example is a local order count $q_t$, interpreted as best-quote depth or as the number of outstanding orders in a price cell. There are two events:

$$q \to q+1, \qquad \text{limit-order arrival at rate } \lambda, \tag{62}$$

$$q \to q-1, \qquad \text{cancellation or removal at rate } \mu q. \tag{63}$$

The model is deliberately small. It is not meant to describe a full order book. It is meant to show how much econometric structure is already contained in one event generator.

## A.1 Master Equation

Let $P(q,t) = \Pr(q_t = q)$. For $q \geq 0$, with $P(-1,t) = 0$, the master equation is

$$\partial_t P(q,t) = \lambda P(q-1,t) + \mu(q+1)P(q+1,t) - (\lambda + \mu q)P(q,t). \tag{64}$$

This equation is the process law before any diffusion approximation is introduced.

## A.2 Doi–Peliti Operator

Let $a^\dagger$ and $a$ denote creation and annihilation operators. The Doi–Peliti generator is

$$\hat{L} = \lambda(a^\dagger - 1) + \mu(1 - a^\dagger)a. \tag{65}$$

The first term creates one order. The second term removes one order with intensity proportional to the current count.

## A.3 Coherent-State Hamiltonian

The coherent-state Hamiltonian is

$$H(q,\varpi) = \lambda(\mathrm{e}^{\varpi} - 1) + \mu q(\mathrm{e}^{-\varpi} - 1). \tag{66}$$

Expanding at $\varpi = 0$ gives

$$H(q,\varpi) = \varpi(\lambda - \mu q) + \frac{1}{2}\varpi^2(\lambda + \mu q) + \frac{1}{6}\varpi^3(\lambda - \mu q) + \frac{1}{24}\varpi^4(\lambda + \mu q) + O(\varpi^5). \tag{67}$$

## A.4 Characteristic Generator

Evaluating the Hamiltonian at imaginary momentum gives the conditional characteristic exponent for a local depth increment:

$$H(q, \mathrm{i}u) = \lambda(\mathrm{e}^{\mathrm{i}u} - 1) + \mu q(\mathrm{e}^{-\mathrm{i}u} - 1). \tag{68}$$

For a short interval $\Delta$, the frozen-state conditional characteristic function is

$$\phi(u \mid q, \Delta) = \exp\left\{\Delta\left[\lambda(\mathrm{e}^{\mathrm{i}u} - 1) + \mu q(\mathrm{e}^{-\mathrm{i}u} - 1)\right]\right\}. \tag{69}$$

This is the object that the demand–supply simulation evaluates with empirical characteristic moments. The cumulants and diffusion approximation below are local derivatives and truncations of the same expression.

## A.5 Diffusion Approximation and Conditional Moments

Keeping only the first two terms in (67) gives

$$\mathrm{d}q_t = (\lambda - \mu q_t)\mathrm{d}t + \sqrt{\lambda + \mu q_t}\mathrm{d}W_t. \tag{70}$$

At sampling interval $\Delta$,

$$\mathbb{E}[\Delta q_t \mid q_t] \approx (\lambda - \mu q_t)\Delta, \tag{71}$$

$$\mathrm{Var}(\Delta q_t \mid q_t) \approx (\lambda + \mu q_t)\Delta. \tag{72}$$

The mean equation is a first-order autoregressive mean-reversion projection:

$$q_{t+\Delta} \approx \lambda\Delta + (1 - \mu\Delta)q_t + \varepsilon_{t+\Delta}. \tag{73}$$

The same generator also implies state-dependent conditional variance.

The higher conditional cumulants follow from the full Hamiltonian:

$$C_k(q) = \left.\frac{\partial^k H(q, \varpi)}{\partial \varpi^k}\right|_{\varpi=0} = \lambda + (-1)^k \mu q. \tag{74}$$

Thus $C_1 = C_3 = \lambda - \mu q$ and $C_2 = C_4 = \lambda + \mu q$. Diffusion econometrics uses $C_1$ and $C_2$. The Doi–Peliti representation keeps $C_3, C_4, \ldots$ as part of the same object.

### A.6 Rare-Event Action

Rare depletion is governed by the action

$$S[q, \varpi] = \int_0^T \left[\varpi_t \dot{q}_t - \lambda(\mathrm{e}^{\varpi_t} - 1) - \mu q_t(\mathrm{e}^{-\varpi_t} - 1)\right] \mathrm{d}t. \tag{75}$$

For a fixed velocity $v = \dot{q}$, the optimized momentum solves

$$v = \lambda \mathrm{e}^{\varpi} - \mu q \mathrm{e}^{-\varpi}. \tag{76}$$

Writing $y = \mathrm{e}^{\varpi}$, the positive solution is

$$y(q, v) = \frac{v + \sqrt{v^2 + 4\lambda\mu q}}{2\lambda}. \tag{77}$$

The local rate function is therefore

$$L(q, v) = v \log y(q, v) - \lambda(y(q, v) - 1) - \mu q(y(q, v)^{-1} - 1). \tag{78}$$

This expression is not the Gaussian quadratic loss. It uses the full jump geometry.

## B Reproducible Numerical Checks

The numerical design has three layers. The first layer is the simulation diagnostic in Sections 8, 9, and 10: event simulation, intensity estimation, regularized ECF estimation, validation, benchmark comparisons, observable distribution forecasts, characteristic diagnostics, bootstrap uncertainty, rare-event diagnostics, identification logic, and input-ready data construction. The second layer is the deliberately simple birth–death derivation used to keep the Doi–Peliti mechanics auditable. The third layer is input readiness. The build records whether valid event-history or transition-panel inputs are available and reports benchmark tables only when normalized histories exist. High-frequency order messages and lower-frequency processor panels are examples of possible input layers, not empirical claims the paper currently settles.

Figure 11: Simulated order depth paths from the birth–death event generator.

The figure shows repeated simulated depth paths from the same birth–death generator. Thin lines are independent event histories, the accent line is one realized path, and the dashed horizontal guide is the stationary mean implied by the arrival and removal rates.

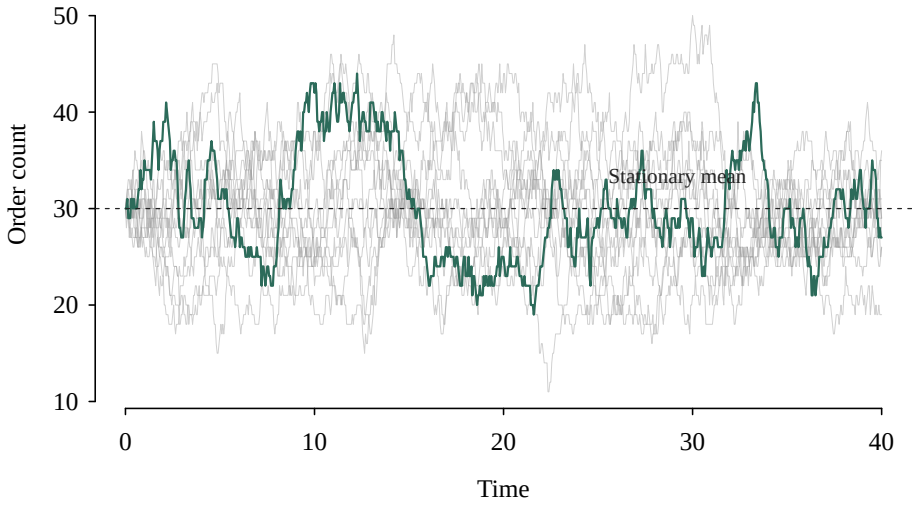



Figure 12: Conditional cumulant rates implied by the same generator.

The plot displays local cumulant rates as functions of the current order count. It shows that the same jump generator delivers drift, variance, and higher cumulants, while a diffusion approximation keeps only the first two.

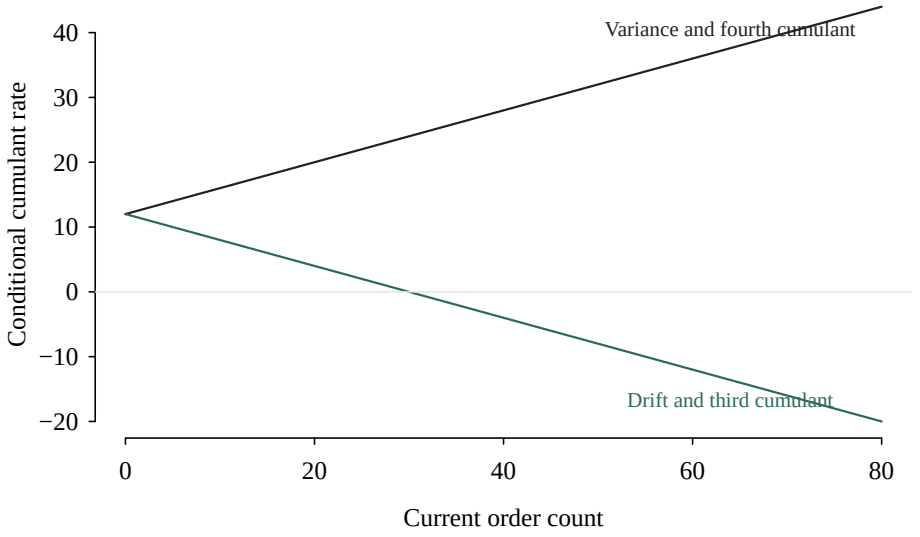
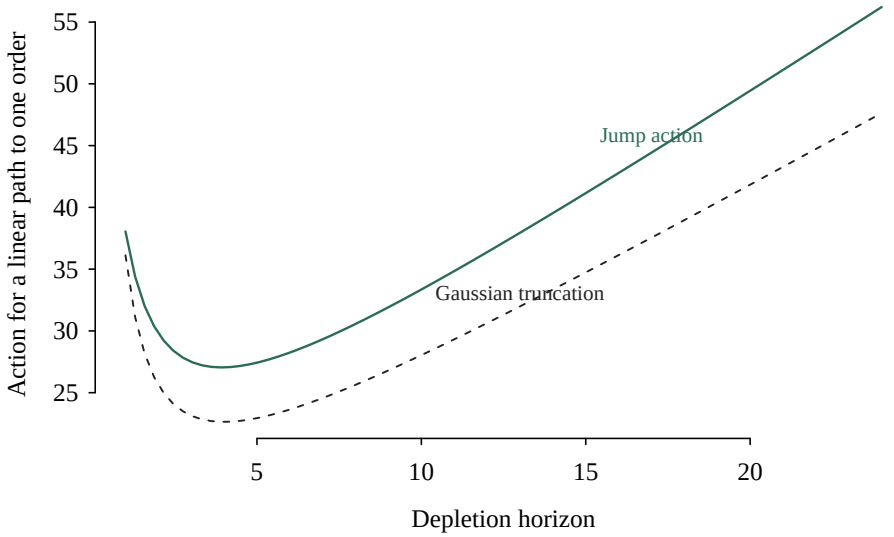


Figure 13: Jump action and Gaussian truncation for a simple liquidity-depletion path.
The graph compares the large-deviation cost of a fixed depletion path under the full jump action and under a Gaussian truncation. The horizontal axis varies the path horizon, so the figure shows when the two approximations assign similar or different rare-event costs.




For the liquidity example, let $q_t$ be the outstanding depth, $T$ the observation span, and $\mathcal{N}_+$ and $\mathcal{N}_-$ the sets of arrival and removal times. Let $N_+ = |\mathcal{N}_+|$ and $N_- = |\mathcal{N}_-|$. If $q_i$ denotes the depth immediately before removal $i$, the exact generator likelihood separates into an arrival part and a cancellation-exposure part:

$$\ell(\lambda, \mu) = N_+ \log \lambda - \lambda T + N_- \log \mu + \sum_{i \in \mathcal{N}_-} \log q_i - \mu \int_0^T q_t \mathrm{d}t. \tag{79}$$

The maximum likelihood estimators are

$$\hat{\lambda} = \frac{N_+}{T}, \qquad \hat{\mu} = \frac{N_-}{\int_0^T q_t \mathrm{d}t}. \tag{80}$$

Second, if only regularly sampled aggregates are used, the diffusion projection estimates

$$\frac{\Delta q_t}{\Delta} = \lambda - \mu q_t + \varepsilon_{t+\Delta}. \tag{81}$$

Here $\varepsilon_{t+\Delta}$ is the projection residual at the sampled horizon. This estimator is intentionally weaker because it discards the event layer. It is included to show how the process-first projection relates

to the generator-first likelihood.

Table 66: Parameter estimates from the reproducible simulation.

The table compares the event-likelihood estimates, the diffusion projection, and the true simulation values for the arrival and removal rates in the birth–death example.

| Method | Arrival rate | Removal rate | Notes |
|---|---|---|---|
| Exact event likelihood | 12.370 | 0.398 | Counts arrivals and integrated cancellation exposure |
| Diffusion drift projection | 12.212 | 0.406 | Least-squares projection of grid increments on lagged depth |
| True simulation value | 12.000 | 0.400 | Used by the reproducible simulation |

Table 67: Symbolic checks for the birth–death derivation.

The table records the main algebraic objects in the birth–death derivation and the corresponding checks used to verify the gain-loss equation, operator form, Hamiltonian, and cumulant hierarchy.

| Object | Formula | Check |
|---|---|---|
| Master equation | $\partial_t P_q = \lambda P_{q-1} + \mu(q+1)P_{q+1} - (\lambda+\mu q)P_q$ | birth and death gain-loss terms |
| Operator | $\hat{L} = \lambda(a^\dagger - 1) + \mu(1 - a^\dagger)a$ | creation for arrivals, annihilation for removals |
| Hamiltonian | $H(q, \varpi) = \lambda(e^{\varpi} - 1) + \mu q(e^{-\varpi} - 1)$ | coherent-state symbol of the generator |
| First cumulant | $\lambda - \mu q$ | drift rate |
| Second cumulant | $\lambda + \mu q$ | diffusion coefficient |
| Higher cumulants | $C_k(q) = \lambda + (-1)^k \mu q$ | full hierarchy retained by the jump generator |

The numerical evidence is intentionally compact. It regenerates the simulation layers, estimates, benchmark tables, figures, observable distribution scores, quantile-regression comparisons, and symbolic checks as a single reproducible record. The input-readiness checks distinguish missing data from failed estimates and report likelihood and ECF losses only when a valid normalized history is available. The purpose is not to substitute simulation for domain evidence, but to make the estimation and evaluation design executable and auditable before a separate application is chosen.

# C Notation and Calculation Map

This appendix collects the notation used in the main calculation chain. It does not introduce a new estimator or a new empirical claim. Its purpose is to make the paper auditable: every reported number should be traceable to an estimand, an estimator, an observable forecast object, and an evaluation rule.

Tables 68 and 69 give the compact map. The first table starts from the adjustment law and records the characteristic-function estimator. The second table records the observable forecast distribution, the scores derived from it, and the uncertainty conventions. Tables 70 and 71 then close the audit loop by matching the reported result-table columns to their defining calculation, sample block, scale convention, and preferred direction. The main text defines the corresponding formulas where they are first used; the tables are a reader's index for those definitions.

The paper separates reporting notation from proof-local notation. Symbols in the calculation maps are the objects that enter reported estimates, forecast distributions, scores, benchmark gaps, and uncertainty summaries. Symbols introduced only inside an assumption, proposition, or proof are local devices for identification, rate, or approximation arguments unless they also appear in one of the notation tables. This convention keeps the displayed calculations auditable without turning the paper into a complete alphabetical glossary.

Two conventions apply throughout the paper. First, all reported losses are computed on a held-out block after the generator, ECF design, regularization, observable mapping, and benchmark library have been selected. Validation can select a specification, but final evidence comes only from the final test block. Second, positive benchmark gaps favor the generator-implied observable distribution: for negatively oriented losses the gap is benchmark loss minus generator loss, while for positively oriented ranking scores it is generator score minus benchmark score. These sign conventions are used in the simulation tables, bootstrap intervals, and benchmark interpretations.

## References


Albrecht, B. C. and Traina, J. (2026). Abduction and the demand curve. Working paper.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, 10(4):1100–1120.

Baez, J. C. and Biamonte, J. (2018). *Quantum Techniques for Stochastic Mechanics*. World Scientific, Singapore. arXiv:1209.3632.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *The RAND Journal of Economics*, 25(2):242–262.

Berry, S. T. and Haile, P. A. (2014). Identification in differentiated products markets using market

Table 68: Calculation map for the adjustment law and ECF estimator.

The table links the adjustment-law estimand and the ECF estimator to the symbols used in the paper, the calculation that produces each object, and the evidence it supports. Panel headers separate the adjustment law from characteristic-function estimation.

| Object | Main symbols | Calculation | Evidence supported |
|---|---|---|---|
| *Adjustment law* | | | |
| State and observable | $X_t$, $Y_t$, $\chi$, $h$ | The state $X_t$ conditions adjustment; a direct mapping $Y_t = \chi(X_t)$ or a measurement law maps adjusted states into outcomes over horizon $h$. | Forecast target and information set. |
| Event generator | $r$, $\nu_r$, $\lambda_r(x)$, $\mathcal{L}_\theta$ | Event $r$ changes the state by $\nu_r$ and arrives at intensity $\lambda_r(x)$; the generator maps test functions into local conditional changes. | Adjustment-law estimand. |
| Transition law | $K_{\theta,\Delta}(\mathrm{d}x' \mid x)$, $\Delta X_i$ | The transition kernel gives the distribution of the future state after horizon $\Delta$; observed increments are compared with this law. | Transition-panel estimand. |
| Likelihood anchor | $\hat{\theta}_{\mathrm{A}}$, $\ell_r$, $Q_T$ | Event histories use point-process likelihood; transition panels use transition likelihood or a prespecified transition fit. | Initialization, benchmark, and regularization anchor. |
| *Characteristic-function estimation* | | | |
| Characteristic law | $H_\theta(x, \mathrm{i}u)$, $\phi_\theta(u \mid x, \Delta)$ | The generator implies a conditional characteristic exponent and a conditional characteristic function for state increments. | Distributional adjustment-law restriction. |
| ECF moments | $\mathcal{U}$, $\mathcal{Z}$, $g_n(\theta)$, $z_j(x)$ | Realized complex exponentials are compared with generator-implied characteristic functions, instrumented by state functions. | Conditional ECF minimum-distance target. |
| Weighting and regularization | $W_n$, $\hat{\Gamma}_n$, $\rho$, $\Omega$, $p_\psi$, $p_\theta$ | Identity weighting gives the first step; a regularized inverse block covariance gives the second step; a validation-selected anchor penalty controls movement from the likelihood or transition fit. | Regularized ECF estimator. |
| Adaptive design | $d$, $\mathcal{D}$, $\mathcal{V}$, $\xi$ | Validation selects frequencies, instruments, block length, weighting rule, and regularization, with an optional complexity price. | Controlled specification search before final testing. |
| Held-out characteristic loss | $L_{\mathrm{CCF}}$, $D_{\mathrm{oracle}}$ | Squared characteristic residuals are averaged over instruments and frequency bands; oracle distance is reported only in simulations. | Adjustment-law distributional fit. |

Table 69: Calculation map for observable forecasts and evaluation.

The table links the generator-implied observable forecast distribution to the reported forecast scores, benchmark gaps, and uncertainty summaries. Panel headers separate forecast construction from evaluation and uncertainty.

| Object | Main symbols | Calculation | Evidence supported |
|---|---|---|---|
| *Observable forecast distribution* | | | |
| Measurement law | $M_h(\mathrm{d}y \mid x, x')$, $F_{Y,\theta}(\cdot \mid x)$ | The transition kernel and measurement law imply a conditional distribution for the observable increment. | Bridge from generator to observables. |
| Simulated forecast distribution | $\widehat{F}_{G,i}$, $S_{\mathrm{sim}}$, $\tilde{y}_i^{(s)}$ | Forward generator paths and measurement draws form an empirical predictive distribution for each held-out observation. | Generator-implied forecast distribution. |
| Quantiles and intervals | $q_{m,i}(\tau)$, $\ell_{m,i,\alpha}$, $u_{m,i,\alpha}$, $\mathcal{A}$ | Quantiles are extracted from a predictive distribution; central intervals use lower and upper quantiles at the stated coverage level. | Quantile and interval forecast objects. |
| Tail and rare-event probabilities | $c$, $R_i$, $\pi_{m,i}$, $\mathcal{D}_m$ | Tail probabilities use the predictive distribution at a prespecified threshold; rare-event probabilities rank held-out observations by predicted risk. | Tail and stress diagnostics. |
| Observable characteristic law | $\varphi_{m,i}^{Y}(v)$, $L_{\mathrm{OCF}}(m)$, $\mathcal{U}_Y$ | The forecast distribution is transformed at observable frequencies and compared with the realized observable transform using standardized conditioning functions. | Frequency-domain check of observable distribution calibration. |
| *Evaluation and uncertainty* | | | |
| Proper forecast scores | $S$, $L_\tau$, $S_{\alpha,i}$, $L_{\mathrm{tail}}$ | Pinball loss, interval score, tail-probability loss, and full-distribution approximations score observable forecast distributions. | Observable distributional value. |
| Ranking and point scores | $N_1$, $N_0$, $\mathrm{AUC}_m$, $\mathrm{Lift}_m$, $\hat{y}_{m,i}$, $\mathrm{RMSE}_m$ | Rare-state counts define the ranking comparison, highest-risk decile concentration summarizes tail concentration, and point-forecast error summarizes scalar forecasts. | Feature-specific benchmark evidence. |
| Benchmark gaps | $\Delta_S(G,B)$, $\hat{\Delta}_S(G,B)$, $F_{G,i}$, $F_{B,i}$ | Population and final-test score gaps compare the generator forecast with prespecified reduced-form benchmarks. | Falsifiable value claim. |
| Decision value, if specified | $a_G^*$, $a_B^*$, $\ell(a,y)$, $\Delta_\ell(G,B)$ | The forecast distribution is mapped into an action by minimizing a stated loss, and held-out realized losses compare generator and benchmark decisions. | Economic decision value. |
| Bootstrap uncertainty | $B_{\mathrm{boot}}$, $\hat{\Delta}_S^{*(b)}$, bootstrap interval, positive share | Contiguous held-out or first-stage blocks are resampled, and the benchmark gap is recomputed under the same sign convention. | Stability of reported gains. |

Table 70: Trace map for forecast and characteristic measures reported in result tables.

The table maps the recurring forecast, characteristic-function, rare-event, and point-forecast columns in the result tables to the formula or construction that produces them. Panel headers separate characteristic diagnostics from observable-distribution and rare-event measures.

| Reported column | Formula or construction | Block and scale | Preferred value |
|---|---|---|---|
| *Characteristic-function and design diagnostics* | | | |
| Validation characteristic criterion | $\hat{V}(d, \rho)$ in (23); the validation moment criterion plus any prespecified complexity price. | Validation block; unscaled unless the table states otherwise. | Lower |
| Moment count | $p_{\psi,d}$, the number of stacked real characteristic moments in design $d$. | Design attribute fixed before final testing. | Not a score |
| Selected penalty | $\hat{\rho}$ from the finite grid in (22) or (23). | Validation-selected regularization value. | Selected |
| Conditional characteristic loss | $L_{\mathrm{CCF}}$ in (18), using standardized state instruments. | Final test block; scaled by $10^4$ when stated in the caption. | Lower |
| Unconditional characteristic loss | $L_{\mathrm{CCF}}$ with only the intercept instrument. | Final test block; same scale as the conditional loss. | Lower |
| Oracle distance | $D_{\mathrm{oracle}}$ in (19). | Simulation only; final test block; same scale as the characteristic loss when stated. | Lower |
| *Observable-distribution scores* | | | |
| Average pinball loss | $\bar{L}_m$ in (33). | Final test block; average over the reported quantile grid. | Lower |
| Continuous ranked probability score approximation | $2\bar{L}_m$, the equal-weight quantile-grid approximation described after (33). | Final test block; same units as the outcome. | Lower |
| Observable characteristic-function loss | $L_{\mathrm{OCF}}(m)$ in (31). | Final test block; scaled by $10^4$ when stated in the caption. | Lower |
| Tail probability loss | $L_{\mathrm{tail}}(m; c)$ in (35), with threshold $c$ fixed before final testing. | Final test block; probability-loss scale. | Lower |
| Crossing share | $C_m$, the share of forecasts with decreasing adjacent quantiles. | Final test block; share or percent. | Lower |
| Simulation paths | $S_{\mathrm{sim}}$, the number of forward generator paths used to form $\widehat{F}_{G,i}$. | Forecast-construction setting fixed before final testing. | Not a score |
| Coverage | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \mathbb{1}\{\ell_{m,i,\alpha} \le y_i \le u_{m,i,\alpha}\}$. | Final test block; percent or share. | Near nominal |
| Average width | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (u_{m,i,\alpha} - \ell_{m,i,\alpha})$. | Final test block; outcome units. | Lower if calibrated |
| Interval score | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} S_{\alpha,i}(m)$, using (34). | Final test block; outcome units. | Lower |
| *Rare-event, point-forecast, and benchmark measures* | | | |
| Brier score | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (R_i - \pi_{m,i})^2$. | Final test block; probability-loss scale. | Lower |
| Area under receiver-operating-characteristic curve | $\mathrm{AUC}_m$ in (36). | Final test block; ranking probability. | Higher |
| Rare-state rate | $\bar{R} = n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} R_i$. | Final test block; prevalence, not a forecast score. | Not a score |
| Highest-risk decile rate | $\bar{R}_{\mathcal{D}_m} = \lvert\mathcal{D}_m\rvert^{-1} \sum_{i \in \mathcal{D}_m} R_i$. | Final test block; highest predicted-risk decile. | Higher |
| Highest-risk decile lift | $\mathrm{Lift}_m = \bar{R}_{\mathcal{D}_m} / \bar{R}$. | Final test block; concentration ratio. | Higher |
| Highest-risk decile count | $N_{\mathcal{D}_m} = \sum_{i \in \mathcal{D}_m} R_i$. | Final test block; count in the highest-risk decile. | Higher |
| Root mean squared error | $\mathrm{RMSE}_m = \{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (y_i - \hat{y}_{m,i})^2\}^{1/2}$. | Final test block; outcome units. | Lower |
| Benchmark gap | $\hat{\Delta}_S(G, B)$ in (30); for ranking scores use generator score minus benchmark score. | Final test block; bootstrap intervals use the same sign convention. | Positive |

Table 71: Trace map for simulation diagnostics and uncertainty measures.

The table maps the simulation-only recovery checks, economic projection diagnostics, rare-action diagnostics, bootstrap summaries, and Monte Carlo summaries to the calculation used in the paper. Panel headers separate first-stage fit, generator projections, rare-event actions, and uncertainty measures.

| Reported column | Formula or construction | Block and scale | Preferred value |
|---|---|---|---|
| *Event-generator fit* | | | |
| Held-out event log likelihood | $n_{\mathcal{T}}^{-1} \sum_{r \in \mathcal{R}} \ell_{r,\mathcal{T}}(\hat{\theta}_{r,m})$, using (49). | Held-out or final test block; per interval. | Higher |
| True-intensity error | $\{(n_{\mathcal{T}}\lvert\mathcal{R}\rvert)^{-1} \sum_{i \in \mathcal{T}} \sum_{r \in \mathcal{R}} [\hat{\lambda}_{r,m}(x_i) - \lambda_{r,0}(x_i)]^2\}^{1/2}$. | Simulation only; held-out states. | Lower |
| Standardized martingale residual | $\max_r \lvert\hat{Z}_{r,m}\rvert$, with $\hat{Z}_{r,m} = \hat{M}_{r,m}/\sqrt{\hat{V}_{r,m}}$. | Held-out block; event-type maximum. | Closer to zero |
| Residual root mean square | $(\lvert\mathcal{R}\rvert^{-1} \sum_{r \in \mathcal{R}} \hat{Z}_{r,m}^2)^{1/2}$. | Held-out block; across event types. | Lower |
| Residual autocorrelation | $\max_r \lvert\hat{a}_{r,m}\rvert$, the largest lag-one compensator-residual autocorrelation. | Held-out block; event-type maximum. | Closer to zero |
| Selection share | Share of resampling or validation draws in which a basis is selected. | Bootstrap or rolling-validation draws. | Stability |
| *Generator projections and economic diagnostics* | | | |
| Drift error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} [\widehat{A}_{E,m}(x_i) - A_{E,0}(x_i)]^2\}^{1/2}$. | Simulation only; held-out states. | Lower |
| Variance error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} [\widehat{B}_{E,m}(x_i) - B_{E,0}(x_i)]^2\}^{1/2}$. | Simulation only; held-out states. | Lower |
| Projection forecast error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (y_i - \hat{y}_{m,i})^2\}^{1/2}$, with the target specified by the table. | Final test block; target units. | Lower |
| Cumulant increment error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} [\hat{C}_{k,m}^E(x_i)\Delta_i - C_{k,0}^E(x_i)\Delta_i]^2\}^{1/2}$. | Simulation only; held-out interval cumulants. | Lower |
| Standardized skewness | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \lvert C_3^E(x_i)\Delta_i / \{C_2^E(x_i)\Delta_i\}^{3/2}\rvert$. | Held-out intervals; scale-free diagnostic. | Descriptive |
| Excess kurtosis | Held-out median of $C_4^E(x_i)\Delta_i / \{C_2^E(x_i)\Delta_i\}^2$. | Held-out intervals; scale-free diagnostic. | Descriptive |
| Price-response slope error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} [\partial_p \hat{A}_{E,m}(x_i) - \partial_p A_{E,0}(x_i)]^2\}^{1/2}$. | Simulation only; held-out states. | Lower |
| Price-response bias | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} [\partial_p \hat{A}_{E,m}(x_i) - \partial_p A_{E,0}(x_i)]$. | Simulation only; held-out states. | Closer to zero |
| Sign agreement | $n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} \mathbb{1}\{\text{sign}(\partial_p \hat{A}_{E,m}) = \text{sign}(\partial_p A_{E,0})\}$. | Simulation only; held-out states. | Higher |
| *Rare-action and uncertainty diagnostics* | | | |
| Action error | $\{n_{\mathcal{T}}^{-1} \sum_{i \in \mathcal{T}} (\hat{I}_{m,i} - I_{0,i})^2\}^{1/2}$. | Simulation only; fixed path and horizon. | Lower |
| Action ratio | $\text{median}_{i \in \mathcal{T}}(\hat{I}_{m,i}/I_{0,i})$. | Simulation only; fixed path and horizon. | Close to one |
| Rank correlation | $\text{Corr}\{\text{rank}(\hat{I}_{m,i}), \text{rank}(I_{0,i})\}$. | Simulation only; held-out states. | Higher |
| Bootstrap interval | $[q_{0.025}\{\hat{\Delta}_S^{*(b)}\}, q_{0.975}\{\hat{\Delta}_S^{*(b)}\}]$. | Block-bootstrap draws; fixed fitted models unless first-stage resampling is stated. | Excludes zero |
| Positive bootstrap share | $B_{\text{boot}}^{-1} \sum_{b=1}^{B_{\text{boot}}} \mathbb{1}\{\hat{\Delta}_S^{*(b)} > 0\}$. | Block-bootstrap draws; same score-gap sign convention. | Higher |
| Monte Carlo ratio | $\text{median}_j(L_{G,j}/L_{B,j})$. | Independent replications; loss criteria only. | Below one |
| Monte Carlo score difference | $\text{median}_j(S_{G,j} - S_{B,j})$. | Independent replications; likelihood or ranking scores. | Above zero |
| Generator-favored share | $N_{\text{rep}}^{-1} \sum_j \mathbb{1}\{L_{G,j} < L_{B,j}\}$, with the analogous inequality reversed for positive scores. | Independent replications. | Higher |
| Replication count | $N_{\text{rep}}$, the number of independent simulation replications. | Monte Carlo design attribute. | Larger |

level data. *Econometrica*, 82(5):1749–1797.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, 31(3):307–327.

Carrasco, M. and Florens, J.-P. (2000). Generalization of GMM to a continuum of moment conditions. *Econometric Theory*, 16(6):797–834.

Chen, X. (2007). Large sample sieve estimation of semi-nonparametric models. In Heckman, J. J. and Leamer, E. E., editors, *Handbook of Econometrics*, volume 6B, pages 5549–5632. Elsevier, Amsterdam.

Doi, M. (1976). Second quantization representation for classical many-particle system. *Journal of Physics A: Mathematical and General*, 9(9):1465–1477.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. *Econometrica*, 50(4):987–1007.

Feuerverger, A. and Mureika, R. A. (1977). The empirical characteristic function and its applications. *The Annals of Statistics*, 5(1):88–97.

Gneiting, T. and Raftery, A. E. (2007). Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102(477):359–378.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90.

Johansen, S. (1991). Estimation and hypothesis testing of cointegration vectors in gaussian vector autoregressive models. *Econometrica*, 59(6):1551–1580.

Koenker, R. and Bassett, G. (1978). Regression quantiles. *Econometrica*, 46(1):33–50.

Newey, W. K. (1997). Convergence rates and asymptotic normality for series estimators. *Journal of Econometrics*, 79(1):147–168.

Pearl, J. (2009). *Causality: Models, Reasoning, and Inference*. Cambridge University Press, Cambridge, 2 edition.

Peliti, L. (1985). Path integral approach to birth-death processes on a lattice. *Journal de Physique*,

46(9):1469–1483.

Singleton, K. J. (2001). Estimation of affine asset pricing models using the empirical characteristic function. *Journal of Econometrics*, 102(1):111–141.

Weber, M. F. and Frey, E. (2017). Master equations and the theory of stochastic path integrals. *Reports on Progress in Physics*, 80(4):046601.