AALBORG UNIVERSITY

DENMARK

MASTER THESIS

Mathematics-economics

Bid–Ask Spread Forecasting

Modelling and Forecasting Bid–Ask Spreads with Integer-Valued Trawl Processes

27th May 2026

Dept. of Mathematical Sciences
Thomas Manns Vej 23, 3rd floor
DK-9220 Aalborg Ø
http://math.aau.dk

AALBORG UNIVERSITY
STUDENT REPORT

Title:
Bid–Ask Spread Forecasting

Theme:
Modelling and Forecasting Bid–Ask Spreads with Integer-Valued Trawl Processes

Project Period:
Spring Semester 2026

Project Group:
M10-12

Authors:
Chris Møller Nørgaard

Supervisors:
Orimar Sauri Arregui

Copies: 1

Numbered Pages: 53

Date of Completion:
27th May 2026

Abstract:

This thesis studies the modelling and forecasting of high-frequency bid–ask spreads using integer-valued trawl processes. The bid–ask spread is discrete, non-negative, persistent, and often overdispersed, making it a natural application for the IVT framework. Six specifications are compared, combining Poisson and negative-binomial Lévy seeds with exponential, inverse-Gaussian, and Gamma trawl functions. The models are estimated by pairwise composite likelihood and evaluated using in-sample fit diagnostics and expanding-window forecast performance.

The empirical application uses five-second bid–ask spread data for four liquid equities. The results show that the negative-binomial seed improves marginal fit by allowing for overdispersion, while the Gamma trawl generally provides the best match to the empirical autocorrelation structure. Forecasting performance is strongly horizon-dependent: Poisson–INGARCH is most competitive at the 5-second horizon, negative-binomial IVT models perform best around the 1-minute horizon, and persistence becomes difficult to beat at longer horizons. Overall, IVT models provide an interpretable and useful framework for spread dynamics, but richer specifications require careful numerical diagnostics and do not uniformly dominate simpler benchmarks.

Preface

This report was written in the period 02/02/2026 to 27/05/2026 by the group M10-12, attending the tenth semester of Mathematics-Economics at Aalborg University. The primary focus of this report is modelling and forecasting bid-ask spreads through integer-valued trawl processes.

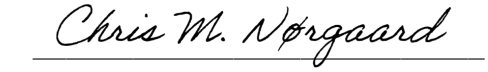
I would like to thank my supervisor, Orimar Sauri Arregui, for his guidance, constructive feedback, and helpful discussions throughout the project.

In this project, the coding and data processing are done in R. The code and data are available on GitHub: `https://github.com/chrismn0/ivt-bid-ask-spread-forecasting.git`.

Generative AI has been used to improve clarity and coherence and to assist with coding tasks.

References follow the APA referencing style and are formatted as

(Author, Year, Chapter or Page number).



Chris Møller Nørgaard
cnor21@student.aau.dk

Nomenclature

Symbol	Description
S_t	Observed bid–ask spread at time t, measured in ticks.
Y_t	Shifted spread process, $Y_t = S_t - 1$.
Y_i^{Δ}	Discretely sampled process, $Y_i^{\Delta} = Y_{i\Delta}$.
Δ	Sampling interval; in the empirical application, $\Delta = 5$ seconds.
q	Forecast horizon measured in five-second steps.
n	Number of observations in an estimation sample.
K	Maximum lag used in the pairwise composite likelihood.
L	Homogeneous Lévy basis.
$L(A)$	Random mass assigned by the Lévy basis to the set A.
L'	Lévy seed of a homogeneous Lévy basis.
A	Trawl set.
A_t	Trawl set shifted to time t.
η	Trawl function defining a monotonic trawl set.
$\mathrm{Leb}(A)$	Lebesgue measure of the set A.
$r(h)$	Autocorrelation function at lag h.
h	Continuous-time lag.
ϑ	Full parameter vector of an IVT model.
$\widehat{\vartheta}_{CL}$	Maximum composite-likelihood estimator.
Θ	Parameter space.
ψ	Parameter vector governing the trawl function.
μ	Poisson Lévy-seed parameter.
m,p	Negative-binomial Lévy-seed parameters.
λ	Exponential-trawl decay parameter.
δ,γ	Inverse-Gaussian trawl parameters.
H,α	Gamma-trawl parameters.
$\ell_{CL}(\vartheta; y)$	Pairwise composite log-likelihood.
$\ell_j(\vartheta)$	Pairwise composite log-likelihood computed using observations up to index j.
$p_{\vartheta}^{(k)}(x,y)$	Model-implied bivariate probability at lag $k\Delta$.
$N_{k,j}(x,y)$	Count of observed lag-k pairs equal to (x,y) up to index j.
$H(\vartheta_0)$	Sensitivity matrix of the composite score.
$V(\vartheta_0)$	Variability matrix of the composite score.
$G(\vartheta_0)$	Godambe information matrix.
CLAIC	Composite-likelihood Akaike information criterion.
CLBIC	Composite-likelihood Bayesian information criterion.
$\widehat{Y}_{i+q\|i}$	Forecast of Y_{i+q}^{Δ} made at forecast origin i.
$\widehat{Y}_{i+q\|i}^{\mathrm{Naive}}$	Naive forecast, $\widehat{Y}_{i+q\|i}^{\mathrm{Naive}} = Y_i^{\Delta}$.
λ_t	Conditional mean in the Poisson–INGARCH benchmark.
$\mathcal{F}_t$	Information available up to time t.
$d_i^{a,b}$	Diebold–Mariano loss differential between models a and b.
ACF	Autocorrelation function.

Symbol	Description
CL	Composite likelihood.
DM	Diebold–Mariano.
GMM	Generalised method of moments.
IG	Inverse Gaussian.
INGARCH	Integer-valued generalized autoregressive conditional heteroskedasticity.
IVT	Integer-valued trawl.
MAE	Mean absolute error.
MSE	Mean squared error.
NB	Negative binomial.
RMSE	Root mean squared error.

Contents

1 | Introduction

Financial liquidity is central to modern financial markets because it affects both trading costs and the ease with which positions can be entered or unwound. One of the most widely used empirical measures of liquidity is the bid–ask spread, defined as the difference between the best quoted ask price and the best quoted bid price. In market microstructure, the spread is not merely a mechanical price gap, but a key object through which trading frictions are analysed. Classic contributions show that bid–ask spreads matter for both transaction costs and asset pricing (Roll, 1984; Amihud and Mendelson, 1986). They are also linked to order-processing costs, inventory considerations, and adverse-selection effects (Stoll, 1989; Huang and Stoll, 1997; Madhavan, 2000). For that reason, modelling the dynamics of the bid–ask spread is of interest not only for short-horizon prediction, but also for understanding how market liquidity evolves over time.

Bid–ask spread data are, however, not straightforward to model. Denoting by (S_t) the bid–ask spread process measured in ticks, the series is integer-valued and observed on a discrete tick grid rather than on a continuous state space. Its trajectories typically display persistence, intraday seasonality, and non-Gaussian marginal behaviour. In such a setting, standard continuous-valued time-series models are not always well suited, while purely discrete-time specifications risk tying the analysis too closely to the chosen sampling frequency. A continuous-time model for integer-valued data is therefore attractive because it respects the tick-based nature of the series while still providing a coherent framework for dependence and forecasting across observation scales.

This motivates the use of integer-valued trawl processes. Introduced by Barndorff-Nielsen et al. (2014), these form a class of stationary, serially correlated continuous-time models for count and integer-valued data. Their main appeal is that they allow the marginal distribution of S_t and the dependence structure of (S_t) to be modelled separately. In the trawl framework, the former is determined by the choice of Lévy seed, while the latter is governed by the trawl function. This separation is especially useful for bid–ask spreads, where one would like to model both the distribution of spread levels and the persistence of dependence in a flexible but tractable way.

A particularly relevant contribution is Bennedsen et al. (2023), who develop likelihood-based inference, model selection, and forecasting methods for continuous-time integer-valued trawl processes, and apply them to financial bid–ask spread data. Their results highlight the importance of modelling both the marginal distribution and the autocorrelation structure carefully. This provides the main methodological point of departure for the present thesis.

The purpose of this thesis is therefore to study whether integer-valued trawl processes provide a useful and empirically credible framework for modelling and forecasting bid–ask spreads. More specifically, the thesis investigates high-frequency quote data for a set of liquid equities and compares alternative specifications for both the Lévy seed and the trawl function. The analysis examines how different assumptions about marginal behaviour and dependence decay affect both in-sample fit and out-of-sample forecast performance. The trawl-based models are evaluated against benchmark forecasting models to assess whether the added modelling flexibility translates into improved predictive performance.

The thesis is guided by the following research question:

> Can integer-valued trawl processes provide a useful framework for modelling and forecasting high-frequency bid–ask spreads, and how do the choices of Lévy seed and trawl function affect empirical fit and forecast performance?

The remainder of the thesis is organised as follows. Chapter 2 introduces the theoretical

framework for trawl processes, with emphasis on the integer-valued setting used in the application. Chapter 3 contains the empirical analysis, beginning with a brief simulation-based discussion of finite-sample considerations and then turning to the modelling and forecasting of bid–ask spread data. Chapter 4 discusses the main findings and limitations of the analysis, and Chapter 5 concludes.

2 | Theoretical framework for trawl processes

This chapter introduces the theoretical framework underlying the trawl-process models used in this thesis. The purpose is not to develop ambit stochastic theory in full generality, but rather to collect the results needed later for modelling bid–ask spreads with integer-valued trawl (IVT) processes.

We begin by introducing the Lévy basis underlying the construction. We then define trawl processes through time-shifted trawl sets and present the properties needed later, including stationarity, dependence induced by set overlap, and the role of the trawl function in shaping the autocorrelation structure. Finally, we outline the estimation and forecasting framework that will be used in the empirical part of the thesis.

Unless otherwise noted, the theoretical framework, notation, and several of the results in this chapter are adapted from Barndorff-Nielsen et al. (2018). The presentation is specialised to the parts needed for integer-valued trawl processes and the empirical application in this thesis. More specialised results for integer-valued trawl processes, inference, and forecasting are cited separately in the relevant sections.

2.1 Lévy basis

We begin by introducing the general notion of a Lévy basis and the associated local characteristics needed for the later construction of trawl processes. Let $(\Omega,\mathcal{F},\mathbb{P})$ denote the underlying probability space, and let $(S,\mathscr{S})$ be a Borel space. We write $\mathcal{B}_b(S)$ for the collection of bounded Borel subsets of S. In the present thesis, and in particular in the application, we will mainly work with $(S,\mathscr{S}) = (\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$.

Definition 2.1 (Lévy basis)
An $\mathbb{R}$-valued *Lévy basis* L on $(S,\mathscr{S})$ is a random measure on $\mathcal{B}_b(S)$ which is *independently scattered* and *infinitely divisible*. Equivalently, L is a family of real-valued random variables $\{L(A) : A \in \mathcal{B}_b(S)\}$ such that:

1. **(Countable additivity)** For any disjoint $A_1, A_2, \ldots \in \mathcal{B}_b(S)$ with $\bigcup_{j=1}^{\infty} A_j \in \mathcal{B}_b(S)$,
$$L\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} L(A_j) \quad \text{a.s.},$$
where the series is assumed to converge almost surely.
2. **(Independent scattering)** If $A_1, A_2, \ldots \in \mathcal{B}_b(S)$ are pairwise disjoint, then $L(A_1), L(A_2), \ldots$ are independent random variables.
3. **(Infinite divisibility)** For every $A \in \mathcal{B}_b(S)$, the law of $L(A)$ is infinitely divisible.

A Lévy basis may be viewed as a set-indexed generalization of a Lévy process. Instead of considering increments over time intervals, one considers random variables $L(A)$ associated with bounded Borel sets $A \in \mathcal{B}_b(S)$. The independently scattered property ensures that

random masses assigned to disjoint sets are independent, while infinite divisibility implies that each $L(A)$ admits a Lévy–Khintchine representation.

Definition 2.2
A Borel measure ν on $\mathbb{R}^d$ is called a *Lévy measure* if

$$\nu(\{0\}) = 0 \quad \text{and} \quad \int_{\mathbb{R}^d} (1 \wedge \|x\|^2)\, \nu(dx) < \infty.$$

The following Lévy–Khintchine representation is standard for Lévy bases; see Barndorff-Nielsen et al. (2018, chapter 5) and Rajput and Rosiński (1989).

Proposition 2.3 (Lévy–Khintchine representation of a Lévy basis)
Let L be a (real-valued) Lévy basis on $(S, \mathscr{S})$. For any $\theta \in \mathbb{R}$ and any $A \in \mathcal{B}_b(S)$, the cumulant (log-characteristic function) of $L(A)$ is given by

$$\begin{aligned} C(\theta; L(A)) &:= \log\Big(\mathbb{E}\Big[e^{\,i\theta L(A)}\Big]\Big) \\ &= i\theta\, \zeta^*(A) - \frac{1}{2}\theta^2 a^*(A) + \int_{\mathbb{R}} \Big(e^{i\theta x} - 1 - i\theta x\, \mathbf{1}_{[-1,1]}(x)\Big)\, n(dx, A). \end{aligned}$$

Here ζ^* is a signed measure on $\mathcal{B}_b(S)$, a^* is a (non-negative) measure on $\mathcal{B}_b(S)$, and $n(\cdot,\cdot)$ is a *generalised Lévy measure* in the following sense: for each fixed $A \in \mathcal{B}_b(S)$, the map $B \mapsto n(B, A)$ defines a Lévy measure on $\mathbb{R}$, while for each fixed Borel set $B \subseteq \mathbb{R}$, the map $A \mapsto n(B, A)$ is a measure on $\mathcal{B}_b(S)$.

The Lévy–Khintchine representation for a Lévy basis is initially formulated in terms of the set functions (ζ^*, a^*, n). To obtain a pointwise or local description, which will later allow us to define the Lévy seed, we introduce a dominating measure and the corresponding Radon–Nikodym derivatives.

Definition 2.4 (Control measure)
Let L be a real-valued Lévy basis on $(S, \mathscr{S})$ with characteristics (ζ^*, a^*, n) as in the Lévy–Khintchine formula. For $B \in \mathcal{B}_b(S)$ define

$$c(B) := |\zeta^*|(B) \;+\; a^*(B) \;+\; \int_{\mathbb{R}} (1 \wedge x^2)\, n(dx, B),$$

where $|\zeta^*|$ denotes the total variation measure of ζ^*. A σ-finite extension of c from $\mathcal{B}_b(S)$ to $\mathscr{S}$ is called a *control measure* for L.

The control measure aggregates the drift, Gaussian variance, and jump activity of the Lévy basis over a set B. Once c is fixed, the law of the basis can be described through local characteristics relative to this measure. Indeed, by construction we have $\zeta^* \ll c$, $a^* \ll c$, and for every Borel set $U \subseteq \mathbb{R}$ the set function $A \mapsto n(U, A)$ is absolutely continuous with respect to c. Hence the Radon–Nikodym derivatives

$$\zeta(z) := \frac{d\zeta^*}{dc}(z), \qquad a(z) := \frac{da^*}{dc}(z), \qquad \ell(U,z) := \frac{d\, n(U,\cdot)}{dc}(z)$$

exist. Then ζ is measurable, a is measurable and non-negative, and for each fixed z the map $U \mapsto \ell(U,z)$ is a measure on $\mathcal{B}(\mathbb{R})$. Equivalently, in kernel notation,

$$\zeta^*(dz) = \zeta(z)c(dz), \qquad a^*(dz) = a(z)c(dz), \qquad n(dx,dz) = \ell(dx,z)c(dz),$$

meaning that for Borel $U \subseteq \mathbb{R}$ and $A \in \mathscr{S}$,

$$n(U,A) = \int_A \ell(U,z)\, c(dz).$$

In typical settings, one may (and we will) choose a version such that for each fixed z, $\ell(\cdot,z)$ is a Lévy measure on $\mathbb{R}$.

Definition 2.5 (Characteristic quadruplet)
The quadruple (ζ, a, ℓ, c) is called a *characteristic quadruplet* of the Lévy basis L if:

1. $\zeta : S \to \mathbb{R}$ and $a : S \to [0,\infty)$ are measurable;
2. for each $z \in S$, $\ell(\cdot,z)$ is a Lévy measure on $\mathbb{R}$, and for each Borel $U \subseteq \mathbb{R}$ the map $z \mapsto \ell(U,z)$ is $\mathscr{S}$-measurable;
3. c is a σ-finite measure on $(S,\mathscr{S})$ such that for all $B \in \mathcal{B}_b(S)$,

$$\int_B |\zeta(z)|\, c(dz) < \infty, \qquad \int_B a(z)\, c(dz) < \infty, \qquad \int_B \int_{\mathbb{R}} (1 \wedge x^2)\, \ell(dx,z)\, c(dz) < \infty.$$

Remark: The choice of the measure c in Definition 2.5 is not uniquely determined by the control measure; changing c typically changes the densities (ζ, a, ℓ), but not the induced law of L as long as the factorisations $\zeta^*(dz) = \zeta(z)c(dz)$, $a^*(dz) = a(z)c(dz)$ and $n(dx,dz) = \ell(dx,z)c(dz)$ represent the same measures. As such, c will mostly be referred to as the *intensity* measure.

With the characteristic quadruplet, the cumulant of $L(A)$ admits the local representation

$$C(\theta; L(A)) = \int_A \left(i\theta\zeta(z) - \frac{1}{2}\theta^2 a(z) + \int_{\mathbb{R}} (e^{i\theta x} - 1 - i\theta x \mathbf{1}_{[-1,1]}(x))\, \ell(dx,z) \right) c(dz).$$

For each $z \in S$, the integrand defines the cumulant of an infinitely divisible law. A random variable with this law is called the *local Lévy seed* at location z, and we denote it by L'_z. Thus,

$$C(\theta; L(A)) = \int_A C(\theta; L'_z)\, c(dz).$$

The local Lévy seed summarizes the drift, Gaussian, and jump behaviour of the Lévy basis at the point z.

2.2 Homogeneous Lévy bases and the Lévy seed

We now consider the homogeneous case. In this setting, the local description of the Lévy basis simplifies substantially, as the local characteristics no longer depend on the location. This leads to the notion of a common Lévy seed.

Definition 2.6 (Factorisable and homogeneous Lévy bases)
Let L be a Lévy basis on $(S, \mathscr{S}) = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$ with characteristic quadruplet (ζ, a, ℓ, c). We call L *factorisable* if the jump kernel $\ell(dx, z)$ does not depend on z. If L is factorisable and, moreover, c is proportional to the Lebesgue measure and the functions $\zeta(z)$ and $a(z)$ are constant in z, then L is called a *homogeneous* Lévy basis. In this case one may write

$$c(dz) = c_0 \operatorname{Leb}(dz) = c_0 \, dz$$

for some constant $c_0 > 0$, where Leb denotes Lebesgue measure on $\mathbb{R}^k$.

Intuitively, homogeneity means that the local characteristics of the Lévy basis do not vary with location. One therefore expects the law of $L(A)$ to be unchanged when the set A is shifted in space, and more generally that the same holds jointly for any finite collection of sets shifted by a common vector.

Proposition 2.7 (Homogeneity and strict stationarity of Lévy bases)
Let L be a real-valued Lévy basis on $(S, \mathscr{S}) = (\mathbb{R}^k, \mathcal{B}(\mathbb{R}^k))$. Then L is *homogeneous* if and only if it is *strictly stationary*, i.e. for every shift $s \in \mathbb{R}^k$ and every finite collection $A_1, \ldots, A_n \in \mathcal{B}_b(S)$ with $A_i + s \in \mathcal{B}_b(S)$,

$$(L(A_1 + s), \ldots, L(A_n + s)) \overset{d}{=} (L(A_1), \ldots, L(A_n)).$$

Proof: We prove the implication from homogeneity to strict stationarity. Let $A_1, \ldots, A_n \in \mathcal{B}_b(\mathbb{R}^k)$ and $\theta_1, \ldots, \theta_n \in \mathbb{R}$. The finite collection $A_1, \ldots, A_n$ generates a finite partition into disjoint sets

$$B_J = \left(\bigcap_{j \in J} A_j \right) \setminus \left(\bigcup_{j \notin J} A_j \right), \qquad \emptyset \neq J \subseteq \{1, \ldots, n\}.$$

Thus

$$A_i = \bigcup_{J \ni i} B_J,$$

and by additivity of the Lévy basis,

$$\sum_{i=1}^{n} \theta_i L(A_i + s) = \sum_{\emptyset \neq J \subseteq \{1, \ldots, n\}} \left(\sum_{i \in J} \theta_i \right) L(B_J + s).$$

Since the sets $B_J + s$ are pairwise disjoint, the corresponding random variables are independent. Hence

$$\log \left(\mathbb{E} \left[\exp \left(i \sum_{i=1}^{n} \theta_i L(A_i + s) \right) \right] \right) = \sum_{\emptyset \neq J \subseteq \{1, \ldots, n\}} C \left(\sum_{i \in J} \theta_i ; L(B_J + s) \right).$$

By homogeneity, $c(dz) = c_0 \operatorname{Leb}(dz)$ and the local characteristics are constant. Therefore

$$C(\theta; L(B_J + s)) = c(B_J + s) C(\theta; L') = c(B_J) C(\theta; L') = C(\theta; L(B_J)).$$

It follows that the joint cumulant function of $(L(A_1+s),\dots,L(A_n+s))$ equals that of $(L(A_1),\dots,L(A_n))$. Hence the joint characteristic functions are identical, and L is strictly stationary.

The converse direction is proved in Barndorff-Nielsen et al. (2018, Proposition 32). In short, strict stationarity implies $C(\theta; L(A+s)) = C(\theta; L(A))$ for all A, s, θ, which by the local Lévy–Khintchine representation forces the control measure to be translation invariant and the local characteristics to be independent of the spatial variable. Hence L is homogeneous. ■

For a homogeneous Lévy basis, the local law is identical at every point. The corresponding local Lévy seed is therefore the same for all z, and we denote this common random variable by L', called the Lévy seed of L. Its cumulant function is

$$C(\theta; L') = i\theta\zeta - \frac{1}{2}\theta^2 a + \int_{\mathbb{R}} \left(e^{i\theta x} - 1 - i\theta x \mathbf{1}_{[-1,1]}(x)\right) \ell(dx).$$

Hence, for every $A \in \mathcal{B}_b(S)$,

$$C(\theta; L(A)) = c(A)\, C(\theta; L').$$

In the remainder of the thesis we work with the normalization $c = \text{Leb}$. Under this convention,

$$C(\theta; L(A)) = \text{Leb}(A) C(\theta; L').$$

This is the key simplification obtained under homogeneity: the distribution of $L(A)$ is determined by the Lebesgue measure of the set and by a single Lévy seed L'. This is precisely the structure that makes the subsequent trawl construction tractable.

2.3 Definition and key properties of trawl processes

We now use the homogeneous Lévy basis from the previous section to define trawl processes. The key idea is to evaluate the basis on a family of time-shifted sets.

Let L be a homogeneous Lévy basis on $(\mathbb{R}^d \times \mathbb{R}, \mathcal{B}(\mathbb{R}^d \times \mathbb{R}))$, and let $A \subset \mathbb{R}^d \times (-\infty, 0]$ be a Borel set with finite Lebesgue measure. The set A is called the *trawl*. For each $t \in \mathbb{R}$, define the shifted trawl

$$A_t := A + (0, t) = \{(\xi, s+t) : (\xi, s) \in A\}. \tag{2.1}$$

Definition 2.8 (Trawl process)
Given a homogeneous Lévy basis L and a trawl set A, the associated *trawl process* is the process $Y = \{Y_t\}_{t\in\mathbb{R}}$ defined by

$$Y_t := L(A_t).$$

Remark: Condition (2.1) ensures that A_t has finite Lebesgue measure for all t. Moreover, since A_t is a translation of A, $\text{Leb}(A_t) = \text{Leb}(A)$. Hence $L(A_t)$ is well-defined for all t.

Proposition 2.9 (Stationarity and cumulant)
Let $Y_t = L(A_t)$ be the trawl process from Definition 2.8. Then Y is strictly stationary and infinitely divisible. Moreover, for every $\theta \in \mathbb{R}$,

$$C(\theta; Y_t) := \log\left(\mathbb{E}\left[e^{i\theta Y_t}\right]\right) = \text{Leb}(A)\, C(\theta; L'). \tag{2.2}$$

Remark: Whenever the cumulants of L' exist, (2.2) implies that the cumulants of Y_t scale linearly with $\mathrm{Leb}(A)$. In particular,

$$\mathbb{E}\left[Y_t\right] = \mathrm{Leb}(A)\ \mathbb{E}\left[L'\right], \qquad \mathrm{Var}\left[Y_t\right] = \mathrm{Leb}(A)\ \mathrm{Var}\left[L'\right].$$

Thus, by selecting the Lévy seed L' one may match any desired infinitely divisible one-dimensional marginal law, while the choice of A governs serial dependence.

2.3.1 The choice of the trawl

Since A_s and A_t can overlap in space-time, part of the same random mass of the Lévy basis can contribute to both variables. This overlap determines the second-order dependence structure of the process.

Proposition 2.10 (Covariance and correlation from trawl overlap)
Let $h \geq 0$. Whenever $\mathrm{Var}\left[L'\right] < \infty$,

$$\mathrm{Cov}\left[Y_t, Y_{t+h}\right] = \mathrm{Leb}(A \cap A_h)\ \mathrm{Var}\left[L'\right].$$

Consequently, the autocorrelation function is

$$r(h) := \mathrm{Corr}\left[Y_t, Y_{t+h}\right] = \frac{\mathrm{Leb}(A \cap A_h)}{\mathrm{Leb}(A)}. \tag{2.3}$$

Proof: By additivity of the Lévy basis,

$$Y_t = L(A_t \setminus A_{t+h}) + L(A_t \cap A_{t+h}),$$

and similarly

$$Y_{t+h} = L(A_{t+h} \setminus A_t) + L(A_t \cap A_{t+h}).$$

The sets

$$A_t \setminus A_{t+h}, \qquad A_{t+h} \setminus A_t, \qquad A_t \cap A_{t+h}$$

are pairwise disjoint. Hence, by independent scattering of the Lévy basis, the corresponding random variables are independent, and we obtain

$$\mathrm{Cov}(Y_t,Y_{t+h}) = \mathrm{Var}\big(L(A_t \cap A_{t+h})\big).$$

Finally, applying Proposition 2.9 to the set $A_t \cap A_{t+h}$ gives

$$\mathrm{Var}(L(A_t \cap A_{t+h})) = \mathrm{Leb}(A \cap A_h)\,\mathrm{Var}[L'],$$

and similarly

$$\mathrm{Var}(Y_t) = \mathrm{Var}(Y_{t+h}) = \mathrm{Leb}(A)\,\mathrm{Var}[L'].$$

Therefore,

$$r(h) = \mathrm{Corr}(Y_t,Y_{t+h}) = \frac{\mathrm{Cov}(Y_t,Y_{t+h})}{\sqrt{\mathrm{Var}(Y_t)\,\mathrm{Var}(Y_{t+h})}} = \frac{\mathrm{Leb}(A \cap A_h)}{\mathrm{Leb}(A)}. \qquad \blacksquare$$

Remark: Equation (2.3) is the key modelling principle: choosing the trawl A amounts to choosing the decay pattern of $r(h)$.

To obtain tractable parametric forms for the autocorrelation function, we now restrict attention to the one-dimensional case, i.e. $S = \mathbb{R} \times \mathbb{R}$ with coordinates (x, s), and to trawls generated by a single deterministic trawl function.

Definition 2.11 (Monotonic trawl and trawl function)
A *monotonic trawl* is defined by

$$A_t = A + (0, t), \qquad A = \{(x, s) : s \leq 0,\ 0 \leq x \leq \eta(s)\} \subset [0,1] \times (-\infty, 0],$$

where the *trawl function* $\eta : (-\infty, 0] \to [0, 1]$ is continuous and non-decreasing, satisfies $\eta(0) = 1$, and

$$\mathrm{Leb}(A) = \int_{-\infty}^{0} \eta(u)\, du < \infty.$$

For a monotonic trawl, the overlap between two shifted trawl sets can be written explicitly. In particular, for $h \geq 0$,

$$\mathrm{Leb}(A \cap A_h) = \int_{-\infty}^{-h} \eta(u)\, du.$$

By Proposition 2.10, this gives

$$r(h) = \frac{\mathrm{Leb}(A \cap A_h)}{\mathrm{Leb}(A)} = \frac{\int_{-\infty}^{-h} \eta(u)\, du}{\int_{-\infty}^{0} \eta(u)\, du}.$$

If η is differentiable, then

$$r'(h) = -\frac{\eta(-h)}{\int_{-\infty}^{0} \eta(u)\, du}.$$

Thus, the tail behaviour of the trawl function determines the rate at which the autocorrelation function decays. In this sense, the choice of η directly specifies the memory of the trawl process. We now consider two standard parametric choices for η.

Exponential trawl. The simplest benchmark is obtained by taking

$$\eta(t) = e^{\lambda t}, \qquad t \leq 0, \quad \lambda > 0.$$

Then

$$\mathrm{Leb}(A) = \int_{-\infty}^{0} e^{\lambda u}\, du = \frac{1}{\lambda}, \qquad \mathrm{Leb}(A \cap A_h) = \frac{e^{-\lambda h}}{\lambda},$$

so the autocorrelation function becomes

$$r(h) = e^{-\lambda h}, \qquad h \geq 0.$$

This specification yields a simple exponentially decaying dependence structure and serves as a natural baseline model.

More flexible decay patterns can be obtained by superposing exponentials. A finite mixture already produces a weighted mixture of exponential decays, while the general continuous superposition takes the form below.

Continuous superposition of exponential trawls. Let π be a probability measure on $(0,\infty)$, and define

$$\eta(t) = \int_0^\infty e^{\lambda t}\,\pi(d\lambda), \qquad t \le 0.$$

If

$$\int_0^\infty \lambda^{-1}\,\pi(d\lambda) < \infty,$$

then $\mathrm{Leb}(A) < \infty$, and the autocorrelation function is

$$r(h) = \frac{\int_0^\infty \lambda^{-1} e^{-\lambda h}\,\pi(d\lambda)}{\int_0^\infty \lambda^{-1}\,\pi(d\lambda)}, \qquad h \ge 0.$$

This representation makes clear that the mixing measure π controls the decay of dependence: placing more mass of π near 0 yields slower decay of $r(h)$, whereas concentrating π away from 0 yields faster decay. Two important special cases are obtained by particular choices of π.

1. **Inverse Gaussian trawl.** A flexible short-memory specification is obtained by choosing π to be an inverse Gaussian distribution. In that case, the resulting superposition trawl has a slower and more flexible decay pattern than the simple exponential benchmark. The corresponding autocorrelation function takes the form

$$r(h) = \exp\left(-\delta\gamma\Big(\sqrt{1+\frac{2h}{\gamma^2}}-1\Big)\right), \qquad h \ge 0.$$

   See Bennedsen et al. (2023, Example 2.4). The details of the calculation are given in the supplementary material to that paper.

2. **Gamma trawl.** Suppose that the mixing measure π has a $\Gamma(1+H,\alpha)$ density, with $H > 0$ and rate parameter $\alpha > 0$,

$$\pi(d\lambda) = \frac{\alpha^{1+H}}{\Gamma(1+H)}\lambda^H e^{-\alpha\lambda}\,d\lambda, \qquad \lambda > 0.$$

   Using the standard Gamma integral,

$$\eta(t) = \int_0^\infty e^{\lambda t}\pi(d\lambda) = \frac{\alpha^{1+H}}{\Gamma(1+H)}\int_0^\infty \lambda^H e^{-(\alpha-t)\lambda}\,d\lambda = \left(1-\frac{t}{\alpha}\right)^{-(H+1)}, \qquad t \le 0.$$

   Applying the overlap formula then gives

$$r(h) = \frac{\int_{-\infty}^{-h}\big(1-\frac{u}{\alpha}\big)^{-(H+1)}\,du}{\int_{-\infty}^{0}\big(1-\frac{u}{\alpha}\big)^{-(H+1)}\,du} = \left(1+\frac{h}{\alpha}\right)^{-H}, \qquad h \ge 0.$$

   In particular, for $H \in (0,1]$, the autocorrelation function is non-integrable, which gives a long-memory specification.

These examples illustrate how different choices of trawl function generate markedly different persistence properties, which will later be relevant when specifying models for bid–ask spread dynamics.

2.3.2 Integer-valued trawl processes (IVT)

Having specified the trawl set and thereby the serial dependence structure, we now restrict attention to the class of trawl processes relevant for our application, where bid–ask spreads are observed in ticks. This leads naturally to integer-valued trawl processes. The idea is to construct a Lévy basis whose values are integer-valued by building it from a marked Poisson random measure. The marks $y \in \mathbb{Z}$ represent jump sizes, while the location variable in $[0,1] \times \mathbb{R}$ determines where these jumps contribute in trawl space. Integrating the marks against the Poisson random measure then yields a random measure with integer-valued increments, which is exactly the structure needed for IVT processes.

Following Bennedsen et al. (2023), let N be a Poisson random measure on

$$\mathbb{Z} \times [0,1] \times \mathbb{R}$$

with intensity measure

$$\nu \otimes \mathrm{Leb} \otimes \mathrm{Leb},$$

where ν is a Lévy measure concentrated on $\mathbb{Z}$, satisfying

$$\sum_{y \in \mathbb{Z}} y^2 \nu(\{y\}) < \infty.$$

For each bounded Borel set $B \subset [0,1] \times \mathbb{R}$, define

$$L(B) := \int_{\mathbb{Z}} y \, N(dy,B) = \sum_{y \in \mathbb{Z}} y \, N(\{y\}, B).$$

Since N is independently scattered, the random measure L is independently scattered as well. Moreover, for each bounded Borel set $B \subset [0,1] \times \mathbb{R}$, the characteristic function of the Poisson random measure yields

$$C(\theta; L(B)) = \log\left(\mathbb{E}\left[e^{i\theta L(B)}\right]\right) = \mathrm{Leb}(B) \sum_{y \in \mathbb{Z}} (e^{i\theta y} - 1)\nu(\{y\}).$$

Hence $L(B)$ is infinitely divisible for every bounded Borel set B, so L is a Lévy basis. Since the cumulant scales linearly with $\mathrm{Leb}(B)$, L is homogeneous with Lévy seed L' satisfying

$$C(\theta; L') = \sum_{y \in \mathbb{Z}} (e^{i\theta y} - 1)\nu(\{y\}).$$

Remark: The above construction fits into the general framework of Definition 2.5. In particular, it yields a factorisable homogeneous Lévy basis on $[0,1] \times \mathbb{R}$ with characteristic quadruplet

$$(\zeta, a, \ell, c), \qquad a = 0, \quad \ell(dx, z) = \nu(dx), \quad c(dz) = \mathrm{Leb}(dz),$$

and drift term

$$\zeta(z) = \int_{[-1,1]} x \, \nu(dx).$$

Hence, the local Lévy–Khintchine representation agrees with the cumulant expression derived above.

The associated trawl process

$$Y_t = L(A_t), \qquad t \in \mathbb{R},$$

is called an *integer-valued trawl (IVT) process*. Although the above construction allows for general integer-valued trawl processes, the empirical application in this thesis concerns bid–ask spreads recorded in ticks. After subtracting the minimum tick, the relevant state space is therefore $\mathbb{N}_0$. For this reason, from this point onward we restrict attention to non-negative IVT processes. In particular, this restriction is satisfied by the Poisson and negative binomial Lévy seeds considered below.

2.3.3 Marginal distributions (Lévy seed choices)

Once the trawl has been specified, the remaining modelling choice for the one-dimensional marginals is the distribution of the Lévy seed L', or equivalently the mark distribution of the underlying Poisson random measure. In this thesis, we focus on two standard choices: the Poisson seed, which yields a parsimonious equidispersed benchmark, and the negative binomial seed, which accommodates overdispersion.

Example (Poisson Lévy seed). Suppose that the Lévy seed is Poisson with mean $\mu > 0$, that is,

$$L' \sim \mathrm{Poi}(\mu).$$

Then its cumulant function is

$$C(\theta; L') = \mu(e^{i\theta} - 1).$$

It follows from (2.2) that

$$Y_t = L(A_t) \sim \mathrm{Poi}(\mu \mathrm{Leb}(A)), \qquad t \in \mathbb{R}.$$

In particular,

$$\mathbb{E}[Y_t] = \mathrm{Var}[Y_t] = \mu \mathrm{Leb}(A).$$

Thus, the Poisson seed yields an equidispersed marginal model and serves as a natural benchmark specification.

Example (Negative binomial Lévy seed). A more flexible alternative is obtained by taking the Lévy seed to be negative binomial. Let

$$L' \sim \mathrm{NB}(m, p),$$

with parameters $m > 0$ and $p \in (0, 1)$, where we use the parametrisation

$$\mathbb{P}(L' = j) = \frac{\Gamma(m + j)}{\Gamma(m)\, j!}(1 - p)^m p^j, \qquad j \in \mathbb{N}_0.$$

Its cumulant function is

$$C(\theta; L') = m\Big(\log(1 - p) - \log\left(1 - pe^{i\theta}\right) \Big).$$

Using (2.2), we obtain

$$Y_t = L(A_t) \sim \mathrm{NB}(m\mathrm{Leb}(A), p), \qquad t \in \mathbb{R}.$$

Hence,

$$\mathbb{E}[Y_t] = \mathrm{Leb}(A)\frac{mp}{1 - p}, \qquad \mathrm{Var}[Y_t] = \mathrm{Leb}(A)\frac{mp}{(1 - p)^2}.$$

In particular,

$$\mathrm{Var}\left[Y_t\right] > \mathbb{E}\left[Y_t\right],$$

so the negative binomial seed allows for overdispersion in the marginal distribution.

The Poisson seed is therefore appropriate as a parsimonious benchmark when the mean and variance are of similar magnitude, whereas the negative binomial seed provides a more flexible alternative when the marginal distribution exhibits overdispersion. This distinction will be relevant in the empirical analysis of bid–ask spreads.

2.4 Estimation and inference for IVT processes

Having specified the dependence structure through the trawl and the one-dimensional marginals through the Lévy seed, we now turn to statistical inference for IVT processes. The aim is to estimate the parameters governing these two components from observed data. A key advantage of the trawl framework is precisely that the serial dependence and the marginal distribution are modelled separately, which provides a natural basis for estimation.

In practice, one may proceed either by moment-based methods or by likelihood-type methods. Moment-based procedures exploit the availability of explicit expressions for low-order moments and the autocorrelation structure, and are useful as simple preliminary estimators. The main inferential approach considered in this thesis, however, is based on pairwise composite likelihood, which remains tractable even when the full likelihood is unavailable.

Generalised method of moments (GMM). A simple moment-based approach is to estimate the trawl parameters by matching the empirical and theoretical autocorrelation functions. If $\hat{r}(h)$ denotes the empirical autocorrelation at lag h, and $r(h;\psi)$ the model-implied autocorrelation for trawl parameter vector ψ, one may estimate ψ by

$$\hat{\psi} \in \arg\min_{\psi} \sum_{h\in H} \left(\hat{r}(h) - r(h;\psi)\right)^2,$$

for a finite set of lags H. Once the trawl parameters have been estimated, the Lévy seed parameters may be obtained by matching low-order sample moments or cumulants using the scaling relation

$$\kappa_i(Y_t) = \mathrm{Leb}(A)\kappa_i(L'),$$

where $\kappa_i(X)$ denotes the ith cumulant of the random variable X. In practice, such moment-based estimators are mainly useful as simple standalone procedures and as starting values for likelihood-based optimisation.

Likelihood-based estimation. Since IVT processes are generally non-Markovian, the full likelihood is typically unavailable except in special cases. Following Bennedsen et al. (2023), we therefore base inference on a pairwise composite likelihood constructed from bivariate distributions. Throughout this subsection, we work with the non-negative IVT specifications introduced above, so that $Y_t \in \mathbb{N}_0$.

Suppose that the continuous-time process Y is observed on an equidistant grid with sampling interval $\Delta > 0$. Define the sampled sequence

$$Y_i^{\Delta} := Y_{i\Delta}, \qquad i = 1,\ldots,n,$$

and denote the observed count values by

$$y_i = Y_i^{\Delta}, \qquad i = 1, \ldots, n,$$

with $y = (y_1, \ldots, y_n)$. For an integer lag $k \geq 1$, corresponding to the time lag $k\Delta$, define

$$CL^{(k)}(\vartheta; y) := \prod_{i=1}^{n-k} \mathbb{P}_\vartheta \left(Y_{i+k}^{\Delta} = y_{i+k}, Y_i^{\Delta} = y_i \right).$$

Here, ϑ denotes the full parameter vector, comprising the trawl parameters and the Lévy-seed parameters, and $\mathbb{P}_\vartheta(B)$ denotes the probability of the event B under parameter value ϑ. The pairwise composite likelihood based on the first K integer lags is then defined by

$$\mathcal{L}_{CL}(\vartheta; y) := \prod_{k=1}^{K} CL^{(k)}(\vartheta; y).$$

Its logarithm is given by

$$\ell_{CL}(\vartheta; y) := \log \left(\mathcal{L}_{CL}(\vartheta; y) \right),$$

and the composite likelihood estimator is

$$\hat{\vartheta}_{CL} \in \arg\max_{\vartheta \in \Theta} \ell_{CL}(\vartheta; y).$$

To evaluate the pairwise probabilities appearing above, we use the trawl-overlap decomposition. For $s < t$,

$$Y_s = L(A \cap A_{t-s}) + L(A \setminus A_{t-s}), \qquad Y_t = L(A \cap A_{t-s}) + L(A_{t-s} \setminus A),$$

where the three components on the right are independent. Hence, for these non-negative IVT specifications, this gives

$$\begin{aligned} \mathbb{P}_\vartheta(Y_s = y_s, Y_t = y_t) = \sum_{x=0}^{\min(y_s, y_t)} & \mathbb{P}_\vartheta\big(L(A \cap A_{t-s}) = x\big) \mathbb{P}_\vartheta\big(L(A_{t-s} \setminus A) = y_t - x\big) \\ & \cdot \mathbb{P}_\vartheta\big(L(A \setminus A_{t-s}) = y_s - x\big). \end{aligned}$$

Here, the sum runs over all possible values of the common overlap component.

As an illustration, consider the Poisson–exponential specification with $L' \sim \mathrm{Poi}(\mu)$ and $\eta(u) = e^{\lambda u}$, $u \leq 0$. For $h = t - s > 0$,

$$\mathrm{Leb}(A \cap A_h) = \frac{e^{-\lambda h}}{\lambda}, \qquad \mathrm{Leb}(A_h \setminus A) = \mathrm{Leb}(A \setminus A_h) = \frac{1 - e^{-\lambda h}}{\lambda}.$$

Hence the component probabilities in the convolution formula are Poisson probabilities:

$$\mathbb{P}_\vartheta(L(A \cap A_h) = x) = p_{\mathrm{Poi}(\mu e^{-\lambda h}/\lambda)}(x),$$

and

$$\mathbb{P}_\vartheta(L(A_h \setminus A) = z) = \mathbb{P}_\vartheta(L(A \setminus A_h) = z) = p_{\mathrm{Poi}(\mu(1-e^{-\lambda h})/\lambda)}(z),$$

where $p_{\mathrm{Poi}(\mu)}(z) = e^{-\mu}\mu^z/z!$. Thus the pairwise likelihood contribution is obtained by inserting these probabilities into the preceding sum, with $h = k\Delta$ for integer lag k. Other seed–trawl combinations are handled analogously.

2.4.1 Inference for the composite likelihood estimator

The pairwise composite likelihood estimator is based on dependent observations from an IVT process, so its large-sample justification requires a weak-dependence condition. Since estimation is carried out on the equidistant grid introduced above, we impose this condition on the sampled sequence

Definition 2.12 (α-mixing for the sampled IVT sequence)
Let $Y_i^\Delta := Y_{i\Delta}$, $i \in \mathbb{Z}$, denote the discretely sampled IVT process. For $q \geq 0$, define

$$\mathcal{F}_{-\infty}^0 := \sigma(Y_i^\Delta : i \leq 0), \qquad \mathcal{F}_q^\infty := \sigma(Y_i^\Delta : i \geq q).$$

The corresponding α-mixing coefficients are

$$\alpha(q) := \sup_{A \in \mathcal{F}_{-\infty}^0,\, B \in \mathcal{F}_q^\infty} |\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|, \qquad q \geq 0.$$

The sampled IVT process is said to be α-mixing if

$$\alpha(q) \to 0 \qquad \text{as } q \to \infty.$$

Thus, $\alpha(q)$ measures the maximal dependence between the σ-fields generated by the distant past and distant future.

For the class of IVT processes considered in this thesis, weak dependence is established under the restriction to monotonic trawls. More precisely, Bennedsen et al. (2023) show that, for an IVT process with monotonic trawl, the mixing coefficients satisfy

$$\alpha(q) = O(r(q)), \qquad h \to \infty,$$

where $r(q)$ denotes the autocorrelation function. Hence the decay of the trawl autocorrelation directly controls the rate at which dependence vanishes.

Under the same framework, and subject to the identification and regularity conditions discussed by Bennedsen et al. (2023), the maximum composite likelihood estimator is consistent as $n \to \infty$ with the sampling interval Δ fixed. Moreover, in the short-memory case, if

$$\lim_{q \to \infty} hr(q) = 0,$$

then

$$\sqrt{n}(\hat{\vartheta}_{CL} - \vartheta_0) \Rightarrow N\Big(0, G(\vartheta_0)^{-1}\Big),$$

where ϑ_0 denotes the true parameter vector of the IVT model and $G(\vartheta_0)^{-1}$ is the inverse Godambe information matrix. The latter has the sandwich form

$$G(\vartheta_0)^{-1} = \mathcal{H}(\vartheta_0)^{-1} V(\vartheta_0) \mathcal{H}(\vartheta_0)^{-1},$$

where $\mathcal{H}(\vartheta_0)$ is the sensitivity matrix and $\mathcal{V}(\vartheta_0)$ is the variability matrix of the composite score. Formally, these may be written as

$$\mathcal{H}(\vartheta_0) = -\lim_{n \to \infty} \frac{1}{n} \mathbb{E}\left[\frac{\partial^2 \ell_{CL}(\vartheta; y)}{\partial \vartheta \partial \vartheta'} \bigg|_{\vartheta = \vartheta_0} \right],$$

and

$$\mathcal{V}(\vartheta_0) = \lim_{n\to\infty} \operatorname{Var}\left(\frac{1}{\sqrt{n}} \frac{\partial \ell_{CL}(\vartheta; y)}{\partial \vartheta}\Big|_{\vartheta=\vartheta_0} \right).$$

The matix $\mathcal{H}$ therefore captures the average curvature of the composite log-likelihood, while $\mathcal{V}$ captures the long-run variation of the composite score. In the present K-lag pairwise likelihood, $\mathcal{V}(\vartheta_0)$ contains the long-run serial covariances between the pairwise score contributions across all included lags; Bennedsen et al. (2023, Theorem 3.3) give the corresponding expanded expression.

Thus, the trawl function matters not only for modelling dependence, but also for the asymptotic behaviour of the estimator: faster decay of $r(h)$ leads to the standard short-memory regime, whereas long-memory trawls may fall outside this standard central limit theorem framework. The Gamma trawl discussed above is an example of such a long-memory specification when $H \in (0,1]$, since its autocorrelation function is then non-integrable.

For feasible inference, $\mathcal{H}(\vartheta_0)$ and $\mathcal{V}(\vartheta_0)$ are replaced by estimates $\hat{\mathcal{H}}(\hat{\vartheta}_{CL})$ and $\hat{\mathcal{V}}(\hat{\vartheta}_{CL})$, giving

$$\hat{G}(\hat{\vartheta}_{CL})^{-1} = \hat{\mathcal{H}}(\hat{\vartheta}_{CL})^{-1} \hat{\mathcal{V}}(\hat{\vartheta}_{CL}) \hat{\mathcal{H}}(\hat{\vartheta}_{CL})^{-1}.$$

While $\hat{\mathcal{H}}$ is straightforward to compute from the curvature of the composite log-likelihood, $\hat{\mathcal{V}}$ is more difficult to estimate in practice. Bennedsen et al. (2023) discuss two approaches in the short-memory case: a kernel-based estimator in the spirit of Newey–West and a simulation-based estimator. Since the present thesis focuses primarily on modelling and forecasting, we do not discuss these variance-estimation procedures in detail.

Information criteria for model selection. Since the pairwise composite likelihood is used in place of the full likelihood, model selection is also based on composite-likelihood analogues of the usual information criteria. In particular, we consider a correction for misspecified likelihoods and define the composite likelihood information criterion

$$\mathrm{CLAIC} = \ell_{CL}(\hat{\vartheta}_{CL}; y) + \operatorname{tr}\Big(\hat{\mathcal{V}}(\hat{\vartheta}_{CL}) \hat{\mathcal{H}}(\hat{\vartheta}_{CL})^{-1} \Big),$$

where $\operatorname{tr}(M)$ denotes the trace of the matrix M. Since the criterion is written in terms of the composite *log*-likelihood, larger values are preferred.

Analogously, a composite-likelihood version of the Schwarz criterion is given by

$$\mathrm{CLBIC} = \ell_{CL}(\hat{\vartheta}_{CL}; y) + \frac{\log(n)}{2} \operatorname{tr}\Big(\hat{\mathcal{V}}(\hat{\vartheta}_{CL}) \hat{\mathcal{H}}(\hat{\vartheta}_{CL})^{-1} \Big),$$

where n is the sample size. Both criteria penalize model complexity through the Godambe-type correction term, with CLBIC imposing the stronger penalty. In the empirical application below, the composite log-likelihood, CLAIC, and CLBIC are used jointly to compare IVT specifications.

2.5 Forecasting integer-valued trawl processes

We now outline how IVT processes can be used for probabilistic forecasting. Let

$$\mathcal{F}_t := \sigma\big((Y_s)_{s\le t}\big)$$

denote the information generated by the process up to time t, and let $h > 0$ be a forecast horizon. The object of interest is the predictive distribution of the future value Y_{t+h}, that is, the conditional distribution of $Y_{t+h} \mid \mathcal{F}_t$.

In general, an IVT process is non-Markovian, so the full conditional distribution $Y_{t+h} \mid \mathcal{F}_t$ is analytically intractable. Following Bennedsen et al. (2023), we therefore adopt a pairwise forecasting approach and approximate the full predictive distribution by the simpler conditional distribution $Y_{t+h} \mid Y_t$. This yields a tractable approximation to the full predictive law, analogous to the use of pairwise composite likelihood in estimation.

Noting the trawl-overlap decomposition from the previous section,

$$Y_t = L(A_t \cap A_{t+h}) + L(A_t \setminus A_{t+h}), \qquad Y_{t+h} = L(A_t \cap A_{t+h}) + L(A_{t+h} \setminus A_t),$$

where the three components on the right are independent, we obtain, for $y_t, y_{t+h} \in \mathbb{N}_0$,

$$\begin{aligned}\mathbb{P}(Y_{t+h} = y_{t+h} \mid Y_t = y_t) = \sum_{l=0}^{\min(y_t, y_{t+h})} & \mathbb{P}(L(A_{t+h} \setminus A_t) = y_{t+h} - l) \\ & \cdot \mathbb{P}(L(A_t \cap A_{t+h}) = l \mid Y_t = y_t)\,.\end{aligned}$$

Here, the sum runs over all possible values of the common overlap component. Once the model parameters are fixed, the distribution of $L(A_{t+h} \setminus A_t)$ is known, so the problem reduces to determining the conditional distribution of the overlap term $L(A_t \cap A_{t+h})$ given Y_t. By Bayes' formula, for $y \in \mathbb{N}_0$ and $l = 0, \ldots, y$,

$$\mathbb{P}(L(A_t \cap A_{t+h}) = l \mid Y_t = y) = \frac{\mathbb{P}(L(A_t \setminus A_{t+h}) = y - l)\, \mathbb{P}(L(A_t \cap A_{t+h}) = l)}{\mathbb{P}(Y_t = y)}.$$

For our two main marginal specifications, the conditional law of the overlap term then takes a particularly convenient form.

Poisson Lévy seed. If the Lévy seed satisfies $L' \sim \mathrm{Poi}(\mu)$, then the two components

$$L(A_t \cap A_{t+h}) \qquad \text{and} \qquad L(A_t \setminus A_{t+h})$$

are independent Poisson random variables with means

$$\mu \,\mathrm{Leb}(A \cap A_h) \qquad \text{and} \qquad \mu \,\mathrm{Leb}(A \setminus A_h),$$

respectively. Since

$$Y_t = L(A_t \cap A_{t+h}) + L(A_t \setminus A_{t+h}),$$

the standard Poisson-splitting property gives

$$L(A_t \cap A_{t+h}) \mid Y_t \sim \mathrm{Bin}\left(Y_t, \frac{\mathrm{Leb}(A \cap A_h)}{\mathrm{Leb}(A)}\right).$$

Equivalently, by Proposition 2.10, the success probability is $r(h)$. The new component satisfies

$$L(A_{t+h} \setminus A_t) \sim \mathrm{Poi}(\mu \,\mathrm{Leb}(A_h \setminus A)).$$

Therefore, inserting the Poisson law of the new component and the binomial conditional law of the overlap component into the convolution formula gives

$$\begin{aligned}\mathbb{P}(Y_{t+h} = y_{t+h} \mid Y_t = y_t) = \sum_{l=0}^{\min(y_t, y_{t+h})} & \frac{(\mu \,\mathrm{Leb}(A_h \setminus A))^{y_{t+h}-l}}{(y_{t+h}-l)!} e^{-\mu \,\mathrm{Leb}(A_h \setminus A)} \\ & \cdot \binom{y_t}{l} r(h)^l (1 - r(h))^{y_t - l}.\end{aligned}$$

Negative binomial Lévy seed. If the Lévy seed satisfies $L' \sim \mathrm{NB}(m,p)$, then the same overlap decomposition applies. In this case, the conditional split

$$\big(L(A_t \setminus A_{t+h}), L(A_t \cap A_{t+h})\big) \mid Y_t$$

has a two-category Dirichlet–multinomial distribution. Equivalently, the overlap component has a beta-binomial distribution. More precisely, for $l = 0, \ldots, Y_t$,

$$\mathbb{P}(L(A_t \cap A_{t+h}) = l \mid Y_t) = \binom{Y_t}{l} \frac{B(l + \alpha_2, Y_t - l + \alpha_1)}{B(\alpha_2,\alpha_1)},$$

where $B(a,b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$ denotes the beta function and

$$\alpha_1 = \mathrm{Leb}(A \setminus A_h)m, \qquad \alpha_2 = \mathrm{Leb}(A \cap A_h)m.$$

Combining this conditional distribution with the marginal negative-binomial law of $L(A_{t+h} \setminus A_t)$ gives the predictive PMF. The resulting expression is available in closed form, but is somewhat more cumbersome than in the Poisson case and is therefore omitted here; see Bennedsen et al. (2023).

In practice, the forecasting formulas depend on the unknown parameter vector through the Lévy seed and the trawl function. These parameters may be estimated using the maximum composite likelihood estimator introduced above. Plugging the estimated parameters into the predictive PMF yields a feasible forecasting procedure for IVT processes. In the empirical part of the thesis, this allows us to construct predictive distributions for future bid–ask spreads and compare the forecasting performance of different trawl specifications.

3 | Empirical analysis

This chapter evaluates the empirical performance of integer-valued trawl models for high-frequency bid–ask spread data. The analysis has two aims. First, it examines how well different IVT specifications reproduce the empirical marginal distribution and serial dependence structure of the spread. Second, it investigates whether the same model classes provide useful out-of-sample forecasts in an expanding-window forecasting design.

The empirical application uses reconstructed limit-order-book data for four liquid NYSE equities sampled at five-second intervals. Since bid–ask spreads are observed in ticks, the data are naturally suited to the integer-valued trawl framework.

The chapter is organised as follows. Section 3.1 describes the data, preprocessing, model specifications and evaluation design. Section 3.2 studies in-sample fit, while Section 3.3 evaluates out-of-sample forecasting performance.

3.1 Data and empirical setup

The empirical application is based on high-frequency bid–ask spread data for Agilent Technologies (A), Discover Financial Services (DFS), Waters Corporation (WAT), and Waste Management (WM). The data are taken from the public replication repository associated with Sauri and Veraart (2022) *Nonparametric estimation of trawl processes* project and are based on reconstructed limit-order-book data sampled at five-second intervals.

Let S_t denote the observed bid–ask spread measured in ticks. Since the minimum spread is one tick, the analysed series is shifted by one tick,

$$Y_t = S_t - 1.$$

This transformation gives a non-negative integer-valued time series while preserving the dynamics of the spread. Forecasts for the original spread can therefore be obtained by adding one tick back.

Following Sauri and Veraart, the first hour of each trading day is discarded to reduce the influence of market-opening effects. The analysis therefore uses observations from 10:30 AM to 4:00 PM. With five-second sampling, this gives $n = 3961$ observations per asset-day. After removing days with data inconsistencies, the final sample contains 458 trading days for A, 452 for DFS, 415 for WAT, and 437 for WM.

For each asset-day, the observations are split chronologically. The first 70%, corresponding to 2772 observations, forms the initial estimation sample and is used for the in-sample diagnostics. The remaining 30%, corresponding to 1189 observations, is reserved for forecast evaluation.

The empirical analysis compares six IVT specifications: Poisson–Exp, Poisson–IG, Poisson–Gamma, NB–Exp, NB–IG, and NB–Gamma. These models vary along two dimensions. The Lévy seed controls the marginal distribution, with the Poisson seed providing a parsimonious benchmark and the negative-binomial seed allowing for overdispersion. The trawl function controls the serial dependence structure, with the exponential trawl providing the simplest specification, while the IG and Gamma trawls allow for more flexible persistence profiles.

All IVT specifications are estimated using the pairwise composite-likelihood framework developed in Chapter 2. Following Bennedsen et al. (2023), the composite likelihood is constructed using pairwise lags up to $K = 10$. This focuses estimation on the short-run

dependence structure, where the empirical autocorrelations are strongest and most precisely estimated. Since the fitted trawl functions imply autocorrelations at all lags, the in-sample diagnostics also assess how well the estimated dependence structure extrapolates beyond the lags used directly in estimation.

3.1.1 In-sample design

The in-sample analysis evaluates the fitted models on the initial 70% estimation window. For each IVT specification, the fitted marginal distribution is compared with the empirical distribution of Y_t, and the fitted autocorrelation function is compared with the empirical autocorrelation function. The models are also compared using composite-likelihood based criteria.

The comparison is reported in two steps. First, a representative trading day for asset A is used to illustrate the fitted marginal distributions and autocorrelation functions in detail. Second, the comparison is repeated across all available asset-days to assess the stability of the model rankings across assets and trading days.

3.1.2 Forecasting design

The out-of-sample analysis uses the remaining 30% of each asset-day. During this holdout period, the estimation window expands as new observations become available. The IVT parameters are updated on a regular one-minute grid, corresponding to every 12 five-second observations.

Let $Y_i^{\Delta} := Y_{i\Delta}$ denote the sampled IVT process, where Δ is the five-second sampling interval. At update index j, the parameter estimate is

$$\hat{\vartheta}_j \in \arg\max_{\vartheta\in\Theta} \ell_j(\vartheta),$$

where $\ell_j(\vartheta)$ is the pairwise composite log-likelihood computed from $Y_1^{\Delta},\ldots,Y_j^{\Delta}$. Between update points, the most recent estimate is kept fixed. Thus, if $\tau(i) \le i$ denotes the most recent update index before or at forecast origin i, the IVT forecast at horizon q is

$$\widehat{Y}_{i+q|i} = \mathbb{E}_{\hat{\vartheta}_{\tau(i)}}\left(Y_{i+q}^{\Delta} \mid Y_i^{\Delta}\right).$$

This ensures that only observations available at the forecast origin are used.

Repeated re-estimation is made feasible by exploiting the discreteness of the spread process. For observations up to index j, the pairwise composite log-likelihood can be written as

$$\ell_j(\vartheta) = \sum_{k=1}^{K}\sum_{i=k+1}^{j} \log p_{\vartheta}^{(k)}\left(Y_{i-k}^{\Delta},Y_i^{\Delta}\right),$$

where

$$p_{\vartheta}^{(k)}(x,y) := \mathbb{P}_{\vartheta}\left(Y_t = x, Y_{t+k\Delta} = y\right)$$

is the bivariate probability implied by the IVT model at lag $k\Delta$. Since many integer-valued pairs are repeated, define

$$N_{k,j}(x,y) := \#\left\{i \in \{k+1,\ldots,j\} : (Y_{i-k}^{\Delta},Y_i^{\Delta}) = (x,y)\right\}.$$

Then the same log-likelihood can be written as

$$\ell_j(\vartheta) = \sum_{k=1}^{K}\sum_{x,y} N_{k,j}(x,y)\log p_{\vartheta}^{(k)}(x,y).$$

This counted representation is algebraically identical to the ordinary pairwise composite likelihood, but avoids repeated evaluation of identical pair contributions. As the estimation window expands, only the counts corresponding to the newly observed lagged pairs need to be updated.

Forecasts are computed at horizons of 5 seconds, 1 minute, 5 minutes and 10 minutes, corresponding to

$$q \in \{1, 12, 60, 120\}$$

five-second steps. Forecast accuracy is measured using squared-error loss, with MSE, RMSE and MAE reported in the output tables. The main relative measure is the MSE ratio with respect to the Naive benchmark,

$$\widehat{Y}^{\text{Naive}}_{i+q|i} = Y^{\Delta}_{i}.$$

A ratio below one indicates an improvement over Naive.

Pairwise forecast differences are assessed using Diebold–Mariano tests based on squared forecast errors; see Diebold and Mariano (1995). For two competing models a and b, the loss differential at forecast origin i is

$$d_i^{a,b} = \left(Y^{\Delta}_{i+q} - \widehat{Y}^{a}_{i+q|i}\right)^2 - \left(Y^{\Delta}_{i+q} - \widehat{Y}^{b}_{i+q|i}\right)^2.$$

The null hypothesis is that the two models have equal expected squared-error loss. Since the forecasts can be overlapping at horizons $q > 1$, the test uses a horizon-dependent long-run variance correction with lag truncation $q - 1$. Because many pairwise comparisons are conducted, p-values are adjusted using the Benjamini–Hochberg procedure (Benjamini and Hochberg, 1995). For the representative-day comparison, the adjustment is applied within each forecast horizon across all pairwise model comparisons. For the multi-day section, day-level tests are computed separately for each trading day, asset, horizon, and model pair, and the p-values are adjusted across trading days within each asset–horizon–model-pair family. Significance is assessed at the 5% level after adjustment.

A Poisson–INGARCH model is also included as a standard observation-driven count time-series benchmark. The model is fitted with the `tscount` package of Liboschik et al. (2017), using the identity-link Poisson INGARCH(1,1) specification

$$\lambda_t = \beta_0 + \beta_1 Y_{t-1} + \alpha_1 \lambda_{t-1},$$

where $Y_t \mid \mathcal{F}_{t-1} \sim \text{Poisson}(\lambda_t)$. The parameters are estimated by conditional maximum likelihood. In the expanding-window forecasting exercise, the INGARCH model is refitted on the same update grid as the IVT models. Multi-step forecasts are computed recursively, replacing future observations by their conditional means.

3.1.3 Finite-sample considerations

Before presenting the empirical results, we briefly summarise a simulation study used to assess the finite-sample behaviour of the estimation procedure in parameter regions relevant for the application. The study is local and application-driven rather than a general investigation of estimator performance over the full admissible parameter space. For each asset and each of the six IVT specifications considered below, the data-generating parameter vector is constructed as the componentwise median of day-level GMM-style calibration vectors computed from the corresponding empirical training windows. For each asset–model scenario, 1000 replications are simulated at sample sizes $n \in \{1000,2500,5000,7500\}$.

The simulation results indicate that parameter recovery is uneven across specifications. The Poisson-based models behave comparatively regularly, with recovery generally improving as the sample size increases. The negative-binomial specifications are more delicate: in particular, the size parameter m can be weakly identified in finite samples and is sometimes estimated at extreme values, even when the optimiser reports convergence. These outlying estimates can dominate mean-based measures such as bias and RMSE.

This instability should not necessarily be interpreted as a complete failure of the negative-binomial models. Rather, it suggests that the richer parametrisations are more vulnerable to finite-sample and numerical identification issues. Consequently, the empirical analysis below places more weight on overall fit and forecast performance than on a literal interpretation of every individual parameter estimate. Detailed recovery plots are reported in Appendix A.1.

3.2 In-sample fit

We now evaluate the in-sample fit of the six IVT specifications using the design described in Section 3.1.1. The analysis first considers a representative trading day for asset A and then turns to the multi-day comparison across all assets.

3.2.1 Representative-day comparison

The representative-day analysis is based on the first available trading day for asset A, corresponding to September 1, 2022. Figure 3.1 shows the intraday spread series after the first hour of trading has been removed. The shaded region marks the first 70% of the observations, which forms the estimation sample used for model fitting and in-sample diagnostics. The remaining 30% of the day is held out for the forecasting exercise considered later.

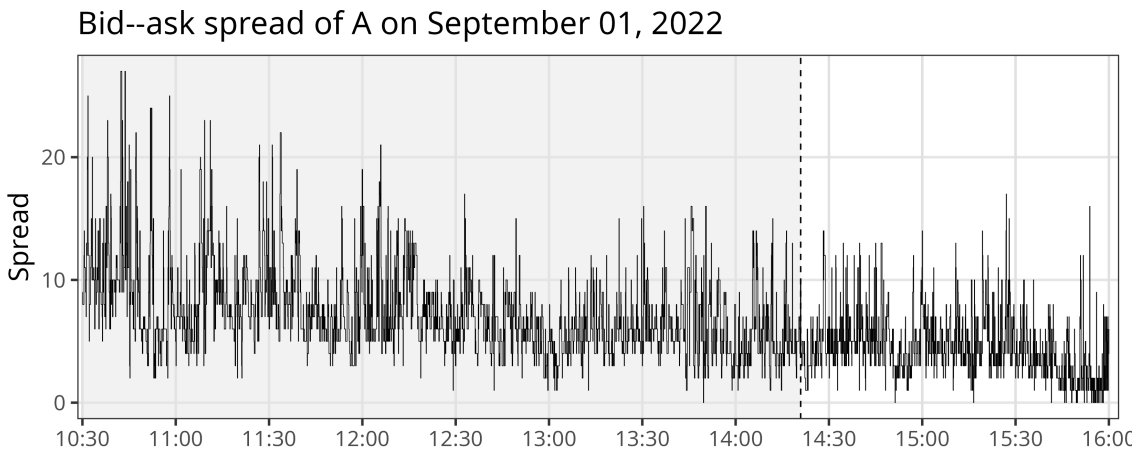


Figure 3.1: Representative-day spread series for asset A on September 1, 2022. The shaded region indicates the first 70% of the day, which is used for model estimation and in-sample diagnostics. The remaining observations are reserved for out-of-sample forecast evaluation.

The figure illustrates several features that motivate the use of flexible integer-valued models. The spread is discrete, non-negative, and exhibits substantial short-run persistence. At the same time, the level of the spread is not constant throughout the day: the series is generally higher and more volatile in the earlier part of the sample and lower towards

the end of the day. This is important when interpreting the in-sample diagnostics, since the IVT specifications considered here are stationary models fitted to a single estimation window.

On this estimation sample, we compare all six IVT specifications: Poisson–Exp, Poisson–IG, Poisson–Gamma, NB–Exp, NB–IG, and NB–Gamma. The comparison focuses on three aspects: the fitted marginal distribution, the fitted autocorrelation function, and the composite-likelihood criteria.

Marginal distribution. Figure 3.2 compares the empirical marginal distribution of the spread on the estimation sample with the fitted model-implied marginal distributions. The clearest distinction is between the Poisson and negative-binomial Lévy seeds. The Poisson specifications produce comparatively concentrated fitted marginals and assign too little probability to parts of the right tail. This is a natural limitation of the Poisson seed, where the variance is tied directly to the mean.

The negative-binomial specifications provide a visibly more flexible marginal fit. They allow for overdispersion and therefore produce heavier-tailed fitted distributions. This improves the fit to the empirical distribution, although the match is not perfect: the negative-binomial models tend to smooth over some of the mass around the empirical modal values. The three negative-binomial specifications are also very close to one another in the marginal comparison, which is expected since they share the same Lévy seed and differ only through the trawl function.

Thus, for the representative day, the marginal fit is driven primarily by the choice of Lévy seed rather than by the choice of trawl function. The negative-binomial seed gives a more credible description of the unconditional spread distribution than the Poisson seed.

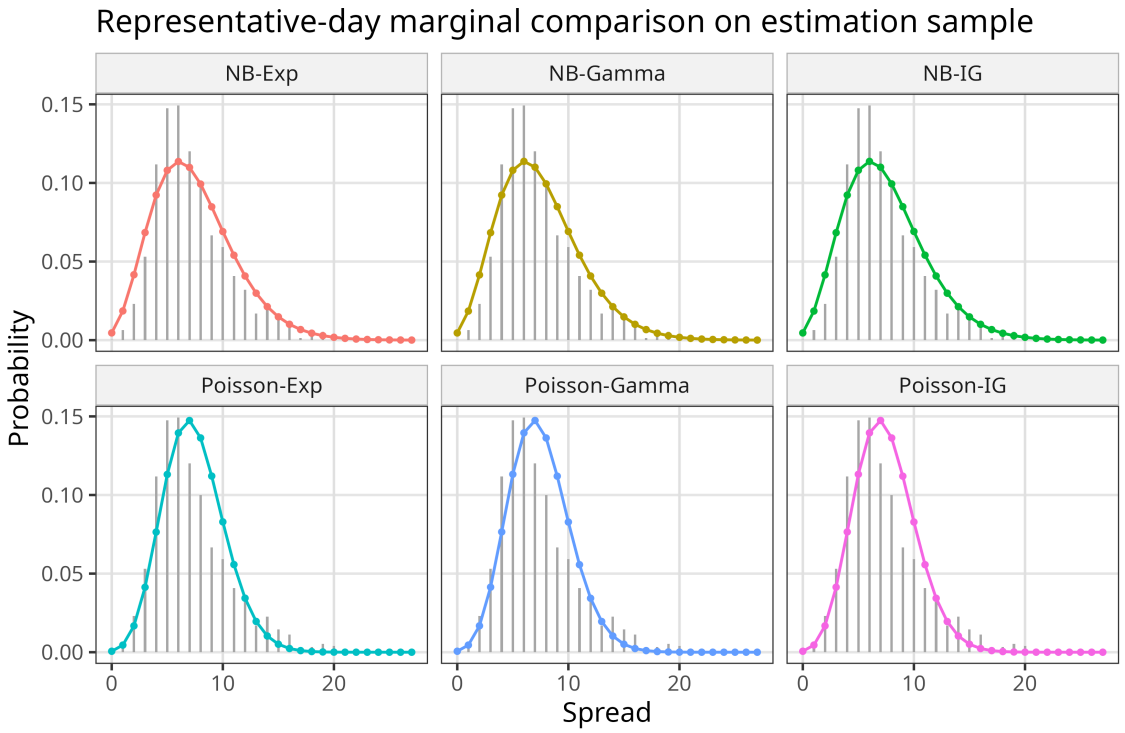


Figure 3.2: Representative-day marginal comparison on the estimation sample. Each panel compares the empirical spread distribution with the fitted marginal distribution of one IVT specification.

Serial dependence. Figure 3.3 compares the empirical autocorrelation function with the fitted ACF implied by each IVT specification. The empirical ACF falls sharply at the shortest lags, but then remains positive over the medium-lag range. This pattern is only partly captured by the fitted trawl functions.

The exponential trawl specifications decay too quickly. This is particularly clear beyond the first few lags, where the fitted exponential ACFs approach zero while the empirical ACF remains positive. The IG and Gamma trawls allow for slower decay and therefore improve the dependence fit relative to the exponential trawl. Among the three trawl families, the Gamma trawl gives the closest visual approximation to the medium-lag persistence, irrespective of whether it is combined with the Poisson or negative-binomial seed.

Nevertheless, none of the six models fully reproduces the empirical dependence structure. Even the more flexible IG and Gamma trawls understate the persistence at longer lags. This discrepancy is important. It suggests that although the IVT framework captures the discreteness and short-run dependence of the spread, a stationary one-factor trawl specification may struggle to reproduce the slower intraday component visible in this representative day. This may reflect genuine long-memory-type dependence, residual intraday seasonality, or local non-stationarity in the spread level.

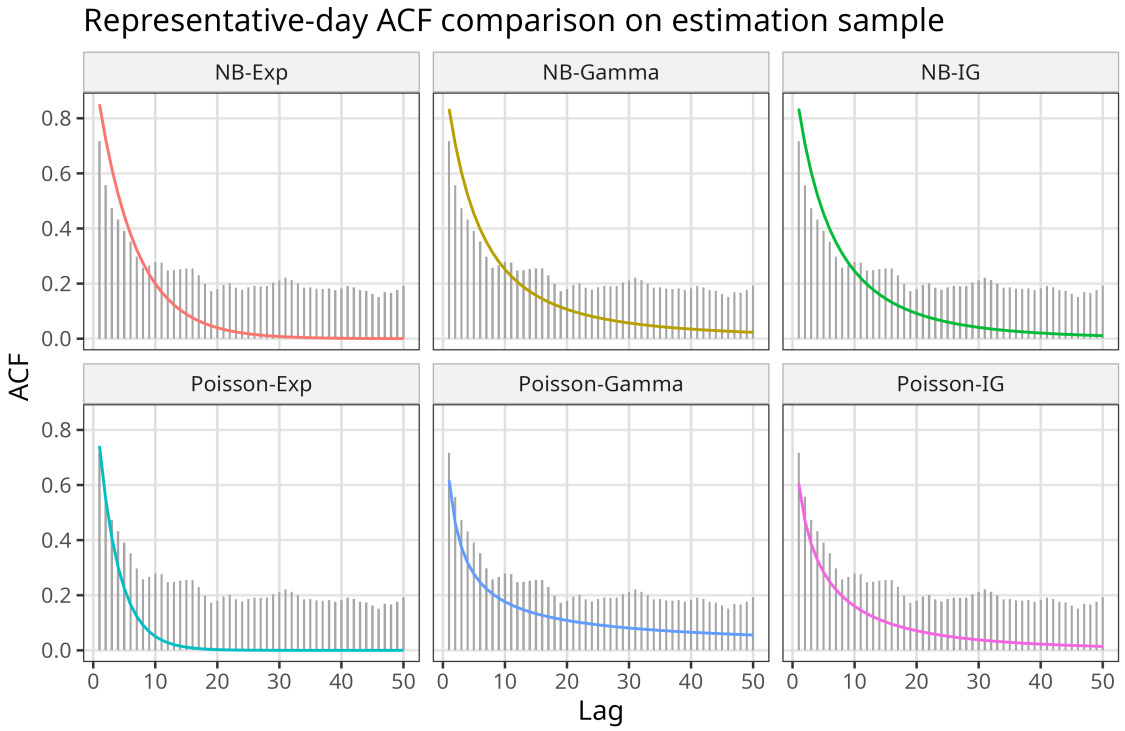


Figure 3.3: Representative-day autocorrelation comparison on the estimation sample. Each panel compares the empirical ACF with the fitted ACF implied by one IVT specification.

Composite-likelihood ranking. The composite-likelihood criteria for the representative day are reported in Table 3.4. The table shows percentage improvements in CL, CLAIC, and CLBIC relative to the Poisson–Exp benchmark, so that positive values indicate better performance than Poisson–Exp. The adjusted criteria are computed using the composite-likelihood Godambe correction.

The ranking is unambiguous. All three negative-binomial specifications dominate all

Table 3.4: Representative-day composite-likelihood performance relative to Poisson–Exp. Positive values indicate percentage improvement relative to the Poisson–Exp benchmark.

Model	ΔCL (%)	ΔCLAIC (%)	ΔCLBIC (%)
NB–Gamma	4.118	4.118	4.118
NB–IG	4.117	4.117	4.116
NB–Exp	4.106	4.106	4.106
Poisson–Gamma	0.259	0.258	0.254
Poisson–IG	0.253	0.253	0.251
Poisson–Exp	0.000	0.000	0.000

three Poisson specifications, and the ordering is the same for CL, CLAIC, and CLBIC. The best model is NB–Gamma, closely followed by NB–IG and NB–Exp. The differences among the negative-binomial specifications are small, while the gap between the negative-binomial and Poisson specifications is substantial. In contrast, changing the trawl function within the Poisson class only gives a modest improvement of about 0.25% relative to Poisson–Exp.

This ranking supports the visual evidence from the marginal comparison. The main improvement in composite-likelihood fit comes from allowing overdispersion through the negative-binomial Lévy seed. The choice of trawl function matters for the shape of the fitted ACF, but it has a smaller effect on the representative-day composite-likelihood ranking than the choice of marginal seed. This also explains why NB–Exp ranks above the Poisson models even though its fitted ACF decays too quickly: the gain in marginal fit is large enough to dominate the overall pairwise likelihood.

Taken together, the representative-day analysis gives two main conclusions. First, the negative-binomial seed is clearly preferable to the Poisson seed for describing the marginal distribution of the spread. Second, the Gamma and IG trawls improve the dependence fit relative to the exponential trawl, but the fitted ACFs still understate the medium-lag persistence observed in the data. The multi-day analysis below examines whether these patterns persist across assets and trading days.

3.2.2 Multi-day in-sample comparison

The representative-day analysis provides a useful illustration of the fitted models, but conclusions based on a single trading day may be sensitive to day-specific features. We therefore repeat the in-sample comparison across all available trading days for each asset. This gives 458 days for A, 415 days for WAT, 452 days for DFS, and 437 days for WM. For each asset-day, each IVT specification is fitted on the first 70% of the observations whenever estimation is successful. The fitted models are then evaluated using the same diagnostics as above: composite likelihood, marginal RMSE, and ACF RMSE up to lag 50.

Figures 3.5 and 3.6 summarize the daily marginal and ACF fit across assets. The marginal RMSE boxplots show that the negative-binomial specifications generally provide a better marginal fit than the Poisson specifications across all four assets. This is consistent with the representative-day comparison and supports the importance of allowing for overdispersion in the spread distribution. The improvement is most pronounced for WAT, DFS, and WM, while the differences are smaller for A. A few large marginal-RMSE outliers are visible, especially for some negative-binomial specifications for A and WAT. Thus, although the negative-binomial seed improves the typical marginal fit, the richer specification can still produce unstable or poor marginal fits on a small number of trading days.

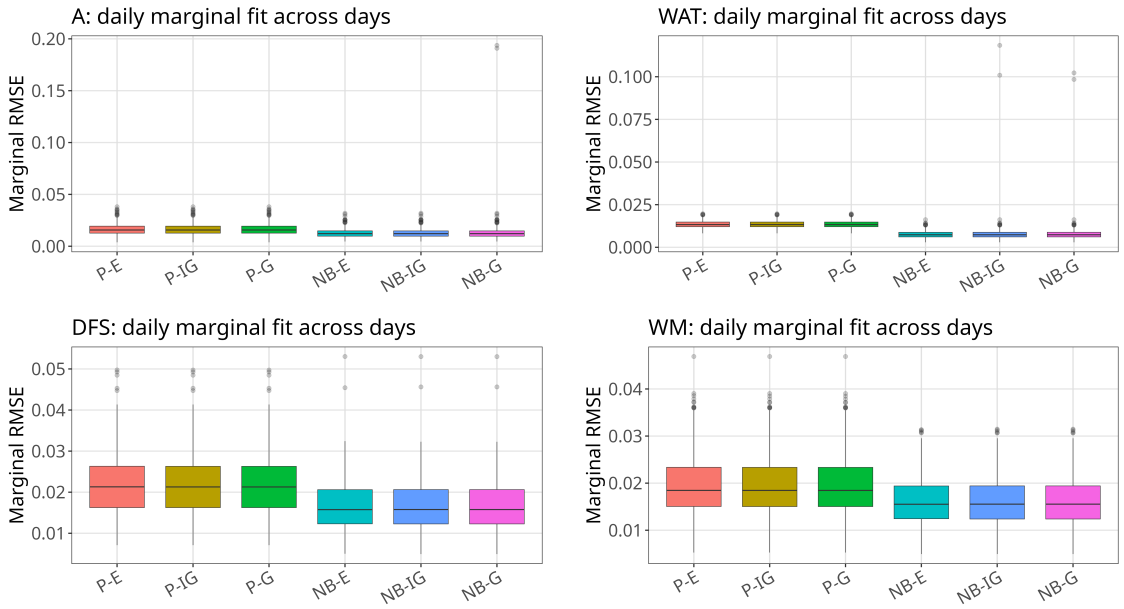


Figure 3.5: Multi-day marginal RMSE across trading days. Each box summarizes the daily marginal fit of one IVT specification for a given asset. Lower values indicate better fit.

Figure 3.6 shows that the exponential trawl specifications generally perform poorly in terms of ACF RMSE. This is consistent with the representative-day results: the exponential trawl imposes a rapidly decaying autocorrelation function and therefore struggles to match the medium-lag persistence observed in the spread data. The IG and Gamma trawls reduce the ACF RMSE substantially, with the Gamma trawl generally giving the lowest and most stable errors across assets. This supports the interpretation that the trawl function is the main driver of dependence fit, and that additional persistence flexibility is useful for modelling bid–ask spreads.

Before reporting winner frequencies, we note that the negative-binomial specifications are not available for every asset-day. Their moment-based initializer requires the empirical variance to exceed the empirical mean; this condition fails on 25 days for A, 62 days for DFS, and 46 days for WM, while WAT is unaffected. These days are retained in the denominator when reporting winner frequencies below.

Table 3.7 reports the two most frequent daily winners under three in-sample criteria: composite likelihood, marginal RMSE, and ACF RMSE. The unadjusted composite-likelihood rankings strongly favour negative-binomial specifications. NB–Gamma is the most frequent CL winner for A, DFS, and WM, while NB–Exp dominates for WAT. This is consistent with the representative-day evidence that overdispersion is important for raw joint fit. The marginal-RMSE winners are more mixed: NB–Exp dominates WAT, whereas Poisson–Gamma and NB–Exp compete closely for the other assets. The ACF-RMSE winners are dominated by Gamma trawl specifications: Poisson–Gamma is the most frequent ACF-RMSE winner for A, WAT, and WM, while NB–Gamma is the most frequent winner for DFS. This confirms that the Gamma trawl provides the most successful dependence fit in the multi-day comparison.

The likelihood ranking in Table 3.7 is based on the unadjusted composite likelihood. We also attempted to compute CLAIC and CLBIC for the full multi-day sample, but the Godambe correction required for these criteria did not complete uniformly across all day-model combinations. The cache export recovered adjusted criteria for 9613 out of

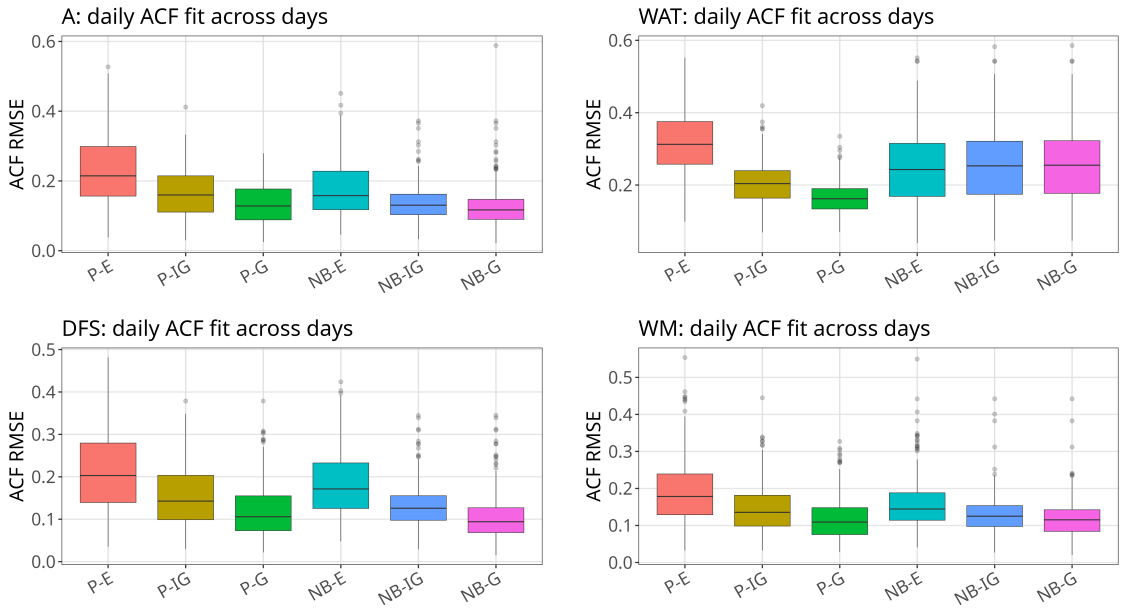


Figure 3.6: Multi-day ACF RMSE across trading days. Each box summarizes the daily ACF fit of one IVT specification for a given asset. Lower values indicate better fit.

Table 3.7: Two most frequent daily winners under in-sample fit criteria. Percentages report the share of all available trading days for the asset on which the model attains the best value of the corresponding criterion. For A, DFS, and WM, negative-binomial fits are unavailable on days where the moment-based initializer requires variance greater than the mean; such days are retained in the denominator.

Asset	Criterion	Most frequent winner	Second most frequent winner
A	CL	NB–Gamma (48.0%)	NB–Exp (44.5%)
A	Marginal RMSE	NB–Exp (32.5%)	Poisson–Gamma (26.9%)
A	ACF RMSE	Poisson–Gamma (44.5%)	NB–Gamma (40.6%)
WAT	CL	NB–Exp (66.7%)	NB–IG (32.3%)
WAT	Marginal RMSE	NB–Exp (69.2%)	NB–Gamma (14.9%)
WAT	ACF RMSE	Poisson–Gamma (70.1%)	NB–Exp (20.5%)
DFS	CL	NB–Gamma (61.9%)	NB–Exp (20.6%)
DFS	Marginal RMSE	Poisson–Gamma (33.4%)	NB–Exp (32.1%)
DFS	ACF RMSE	NB–Gamma (47.1%)	Poisson–Gamma (40.7%)
WM	CL	NB–Gamma (50.8%)	NB–Exp (36.4%)
WM	Marginal RMSE	Poisson–Gamma (37.3%)	NB–Exp (26.8%)
WM	ACF RMSE	Poisson–Gamma (53.8%)	NB–Gamma (33.2%)

10572 fitted day-model rows, corresponding to an overall completion rate of 90.9%. Since the missing adjusted criteria are concentrated in numerically more difficult specifications, especially NB–Gamma, the adjusted criteria are not used as the primary multi-day ranking. Appendix A.2 reports completion rates and adjusted rankings on the subset of days for which all six specifications have finite adjusted criteria. On this reduced common subset, CLAIC and CLBIC favour the simpler Poisson–Exp model. This suggests that the raw likelihood gains from the richer specifications are not always large enough to compensate for their estimated effective complexity. Thus, the unadjusted composite-likelihood rankings should be interpreted as raw in-sample fit comparisons rather than as fully penalized model-selection results.

Overall, the multi-day in-sample evidence supports three conclusions. First, the unadjusted composite-likelihood rankings favour negative-binomial specifications whenever the negative-binomial fit is available, which is consistent with the importance of overdispersion in the spread distribution. However, the adjusted criteria reported in Appendix A.2 show that, after effective complexity adjustment, the simpler Poisson–Exp model is often preferred on the reduced common subset. Second, the best marginal RMSE is not always obtained by the same models that maximize the composite likelihood, showing that the composite likelihood reflects the joint pairwise distribution and not the marginal distribution alone. Third, the ACF RMSE comparison consistently points toward the Gamma trawl family, suggesting that the additional persistence flexibility of the Gamma trawl is useful for matching the serial dependence of bid–ask spreads. The forecasting analysis therefore provides an important complementary assessment of whether these in-sample differences translate into improved predictive performance in an expanding-window forecasting design.

3.3 Forecasting

We now evaluate the out-of-sample forecasting performance of the six IVT specifications using the expanding-window design described in Section 3.1.2. The analysis follows the same two-level structure as the in-sample comparison. We first consider the representative trading day for asset A in order to illustrate the forecast behaviour of the competing models. We then turn to a multi-day comparison across assets, which provides the main evidence on the robustness of the forecasting results.

The IVT forecasts are conditional-mean forecasts implied by the fitted trawl model. Thus, the fitted trawl function enters the forecasting problem through the amount of dependence between the current spread and the future spread. A trawl specification with faster decay places less weight on the current spread and mean-reverts more quickly towards the fitted marginal mean.

3.3.1 Representative-day forecast comparison

We first evaluate the expanding-window forecasts on the representative trading day for asset A. The initial estimation window is the same first 70% of the day used in the in-sample analysis, while the remaining 30% is used for forecast evaluation. Forecasts are computed at horizons of 5 seconds, 1 minute, 5 minutes, and 10 minutes. The IVT models are refitted during the holdout period using the expanding-window design described in Section 3.1.2, with refits performed every 12 observations. For this representative day, all forecast models have full coverage at all horizons, and all expanding-window IVT refits are accepted.

MSE ratios. Figure 3.8 reports MSE ratios relative to the Naive benchmark. Values below one indicate that the model improves on the Naive forecast. The short-horizon results are favourable for the model-based forecasts. At the 5-second horizon, all IVT specifications and the Poisson-INGARCH benchmark improve on Naive in terms of MSE. The best model at this horizon is Poisson-INGARCH, while Poisson–Exp is the best IVT specification. This suggests that at the shortest horizon, the most recent spread contains substantial predictive information, and relatively fast mean reversion can improve squared-error performance.

At the 1-minute horizon, the ranking changes. The best models are NB–Gamma and NB–IG, with MSE ratios of 0.886 and 0.897 relative to Naive. Thus, the negative-binomial seed combined with a more persistent trawl function gives the best representative-day performance at this horizon. At the 5-minute horizon, the best model is Poisson–Gamma, with an MSE ratio of 0.949. The improvement over Naive is smaller than at shorter horizons, but the ranking again points towards the value of the Gamma trawl for medium-horizon forecasts. At the 10-minute horizon, however, the Naive forecast performs best. All model-based forecasts have MSE ratios above one, although Poisson–Gamma remains the closest competitor. Thus, on the representative day, the gains from the IVT forecasts are concentrated at short and intermediate horizons and disappear at the longest horizon considered.

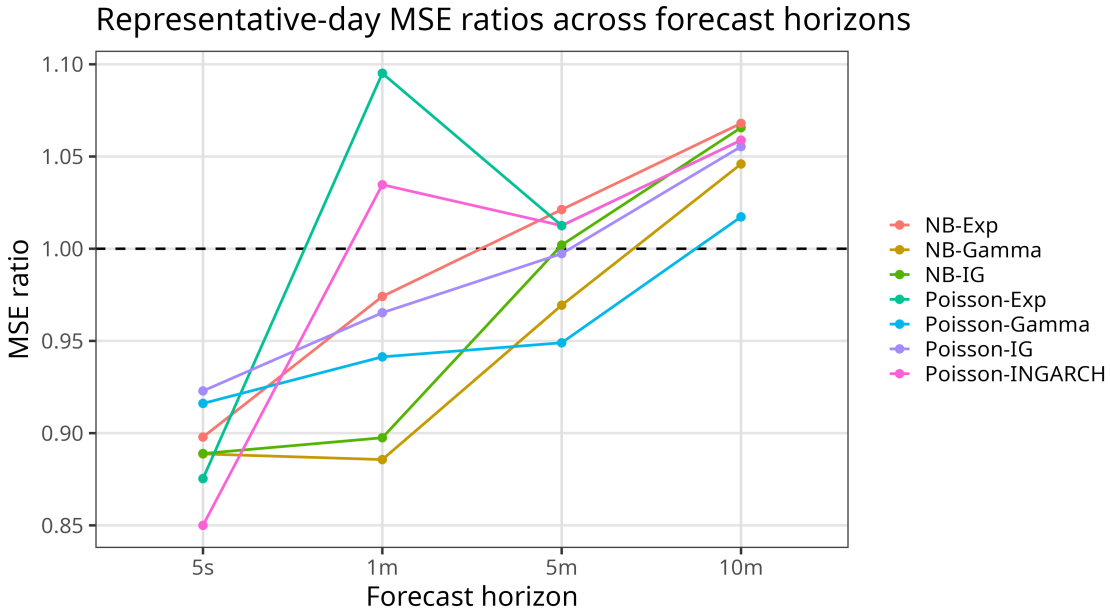


Figure 3.8: Representative-day MSE ratios across forecast horizons for asset A. Ratios are computed relative to the Naive benchmark, so values below one indicate lower MSE than Naive.

Diebold–Mariano tests. Figure 3.9 reports pairwise Diebold–Mariano tests based on squared forecast errors. At the 5-second horizon, the model-based forecasts significantly outperform Naive after multiple-testing adjustment. This supports the visual evidence from the MSE ratios. At longer horizons, the evidence against Naive becomes weaker. Although several IVT specifications have lower MSE than Naive at the 1-minute and 5-minute horizons, the pairwise DM tests do not show statistically significant improvements over Naive after adjustment. At the 10-minute horizon, Naive is generally preferred.

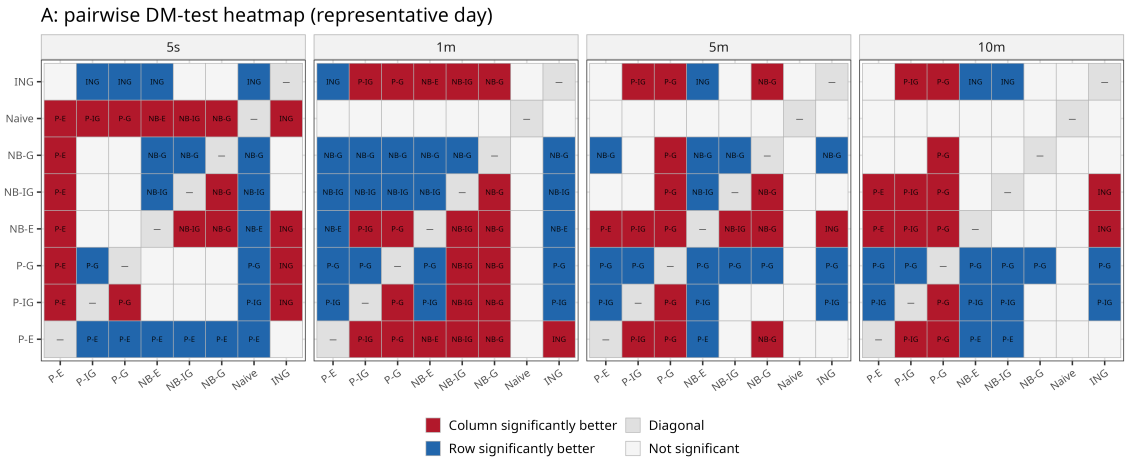


Figure 3.9: Pairwise Diebold–Mariano test heatmap for representative-day forecasts of asset A after Benjamini–Hochberg adjustment. Tests are based on squared forecast errors. Red cells indicate that the column model is significantly better, blue cells indicate that the row model is significantly better, and white cells indicate no significant difference after adjustment.

Overall, the representative-day forecasting results show that the IVT forecasts can improve on simple Naive at short horizons, but that the advantage is horizon-dependent. The strongest gains occur at the 5-second and 1-minute horizons. At 5 minutes the improvement is modest, and at 10 minutes the Naive benchmark performs best. This suggests that the fitted trawl dependence structure contains useful short-run predictive information on this day, but that the benefit of the parametric IVT forecasts weakens as the horizon increases.

3.3.2 Multi-day forecast comparison

The representative-day forecast comparison gives a detailed view of one trading day, but the main forecasting evidence comes from the multi-day analysis. We therefore repeat the expanding-window forecasting exercise for all available trading days and assets. Forecasts are evaluated at horizons of 5 seconds, 1 minute, 5 minutes, and 10 minutes. The main forecast measure is the daily MSE ratio relative to the Naive benchmark. Values below one indicate that a model improves on Naive for the corresponding asset-day and forecast horizon. The full day-level pairwise Diebold–Mariano heatmaps are reported in Appendix A.3; the main text summarizes the results using median MSE ratios and winner frequencies.

Figure 3.10 reports the median daily MSE ratios across forecast horizons for each asset. The results are strongly horizon-dependent. At the 5-second horizon, the Poisson–INGARCH benchmark performs best for A, DFS, and WM, with median MSE ratios of 0.891, 0.853, and 0.860, respectively. This suggests that at the shortest horizon, a simple observation-driven model can exploit very local dependence in the spread effectively. For WAT, the 5-second comparison is less dominated by Poisson–INGARCH; the negative-binomial IVT specifications have the lowest median MSE ratios, although the differences are small.

At the 1-minute horizon, the IVT models become more competitive. The best median MSE ratios are obtained by negative-binomial specifications for all four assets: NB–Gamma

for A, DFS, and WM, and NB–Exp for WAT. The median MSE ratios are around 0.85–0.91, indicating a clear improvement over Naive. This is the horizon at which the IVT models perform most consistently across assets. The result is also consistent with the role of the trawl function in the conditional-mean forecast: at a one-minute horizon, retaining a flexible amount of dependence on the current spread is useful, while Naive begins to overstate the relevance of the most recent observation.

At longer horizons, the gains become weaker. For A, DFS, and WM, the Naive benchmark has the lowest median MSE ratio at both the 5-minute and 10-minute horizons. The best IVT competitors at these horizons are usually Gamma-trawl specifications, especially Poisson–Gamma and NB–Gamma, but their median MSE ratios are generally above one. WAT is the main exception: the negative-binomial IVT models still have median MSE ratios below one at the 5-minute horizon, and NB–Gamma remains slightly below one at the 10-minute horizon. Overall, the evidence suggests that the IVT forecasts are most useful at short and intermediate horizons, while the advantage becomes less robust as the forecast horizon increases.

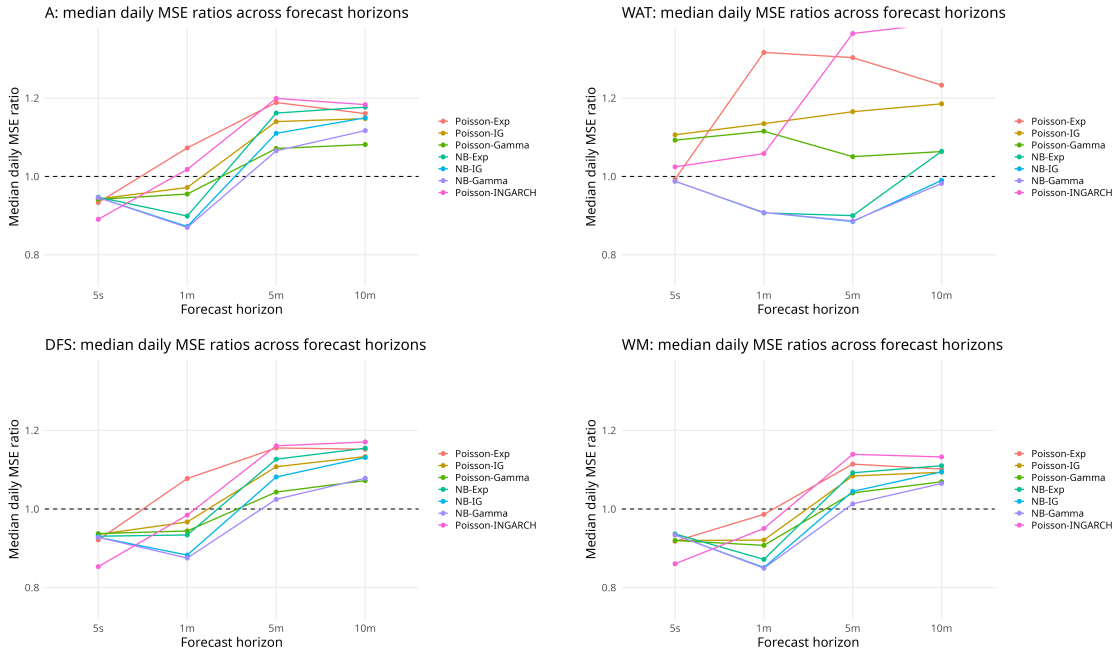


Figure 3.10: Median daily MSE ratios across forecast horizons. Ratios are computed relative to the Naive benchmark, so values below one indicate lower MSE than Naive.

Table 3.11 reports the two most frequent daily winners for each asset and horizon. The winner frequencies reinforce the horizon-dependent pattern from Figure 3.10. At the 5-second horizon, Poisson–INGARCH is the most frequent winner for all four assets, although the dominance is much stronger for A, DFS, and WM than for WAT. At the 1-minute horizon, negative-binomial IVT specifications are the most frequent winners for all assets. NB–Gamma wins most often for A, DFS, and WM, while NB–Exp wins most often for WAT. At the 5-minute and 10-minute horizons, Naive becomes the most frequent winner for all assets, although Poisson–Gamma and NB–Gamma remain important competitors. This shows that the model ranking is not stable across horizons: the best short-horizon model is not necessarily the best medium-horizon model.

The day-level Diebold–Mariano tests, reported in Appendix A.3, provide a complementary view of these results. They show that statistically significant pairwise differences are

Table 3.11: Two most frequent daily forecast winners by asset and horizon. Percentages are computed over valid trading days with finite daily MSE ratios.

Asset	Horizon	Most frequent winner	Second most frequent winner
A	5s	Poisson–INGARCH (80.9%)	Poisson–Exp (5.8%)
A	1m	NB–Gamma (46.0%)	Naive (25.1%)
A	5m	Naive (55.6%)	Poisson–Gamma (26.8%)
A	10m	Naive (58.4%)	Poisson–Gamma (29.9%)
WAT	5s	Poisson–INGARCH (33.3%)	Poisson–Exp (23.4%)
WAT	1m	NB–Exp (48.7%)	NB–IG (16.1%)
WAT	5m	Naive (30.4%)	NB–Gamma (26.3%)
WAT	10m	Naive (44.6%)	Poisson–Gamma (26.5%)
DFS	5s	Poisson–INGARCH (87.4%)	Poisson–Exp (4.3%)
DFS	1m	NB–Gamma (49.3%)	Naive (25.5%)
DFS	5m	Naive (51.9%)	NB–Gamma (25.6%)
DFS	10m	Naive (55.3%)	Poisson–Gamma (27.2%)
WM	5s	Poisson–INGARCH (89.6%)	Poisson–Exp (2.9%)
WM	1m	NB–Gamma (52.9%)	Naive (19.4%)
WM	5m	Naive (49.9%)	Poisson–Gamma (32.2%)
WM	10m	Naive (53.6%)	Poisson–Gamma (34.6%)

most common at the shorter horizons, while many longer-horizon comparisons are either insignificant or split across trading days. This is consistent with the MSE-ratio evidence: model-based forecasts can improve on Naive at short horizons, but the advantage becomes less systematic as the horizon increases.

Taken together, the multi-day forecasting results suggest three main conclusions. First, forecast performance is highly horizon-dependent. Poisson–INGARCH is strongest at the 5-second horizon, negative-binomial IVT models are strongest at the 1-minute horizon, and Naive becomes difficult to beat at 5 and 10 minutes. Second, the negative-binomial IVT specifications provide the most consistent IVT forecasting gains, especially at the 1-minute horizon. Third, the Gamma trawl remains relevant at longer horizons, where Poisson–Gamma and NB–Gamma are usually the closest competitors to Naive. Thus, the in-sample evidence that overdispersion and persistence flexibility matter does translate into forecasting gains, but mainly at short and intermediate horizons rather than uniformly across all forecast horizons.

4 | Discussion

This chapter discusses the main empirical findings and their implications for modelling and forecasting bid–ask spreads with integer-valued trawl processes. The purpose is to interpret what the empirical results imply about the usefulness and limitations of the IVT framework in this application.

The empirical analysis shows that IVT models provide a natural and flexible framework for bid–ask spread data. The spread is integer-valued, non-negative, persistent, and often overdispersed, which makes it well suited to the modelling structure considered in this thesis. At the same time, the results also show that no single specification dominates across all diagnostics, assets, and forecast horizons. The main conclusion is therefore nuanced: IVT models are useful and interpretable, but their practical performance depends strongly on the choice of marginal seed, trawl function, forecast horizon, and evaluation criterion.

4.1 Marginal fit and dependence fit

A central advantage of the IVT framework is that it separates two modelling components. The Lévy seed determines the marginal distribution, while the trawl function determines the serial dependence structure. The empirical results show that this distinction is useful in practice.

For the marginal distribution, the negative-binomial seed generally improves the fit relative to the Poisson seed. This is most clearly visible in the representative-day comparison, where the Poisson specifications are too concentrated and assign too little probability to parts of the right tail. The negative-binomial models provide a better description of the overdispersion in the spread distribution. The multi-day comparison supports the same interpretation, although the strength of the improvement varies across assets and days. Thus, allowing for overdispersion is important for modelling the unconditional distribution of bid–ask spreads.

For the dependence fit, the exponential trawl often decays too quickly and therefore struggles to reproduce the medium-lag persistence observed in the empirical autocorrelation function. The IG and Gamma trawls improve the dependence fit, with the Gamma trawl generally performing best in terms of ACF RMSE. This suggests that the additional persistence flexibility of the Gamma trawl is useful for spread data.

These findings imply that the best model depends on which aspect of the data one wants to emphasize. The negative-binomial seed is important for marginal fit, while the Gamma trawl is important for dependence fit. A model with strong marginal fit is not necessarily the best dependence model, and a model with the best ACF fit is not necessarily the best according to composite likelihood. Thus, model comparison for IVT processes should not rely on a single diagnostic alone.

4.2 Composite likelihood and model complexity

The in-sample likelihood results also show that model complexity matters. The unadjusted composite likelihood generally favours the negative-binomial specifications, especially NB–Gamma and NB–Exp. This is consistent with the marginal evidence, since the negative-binomial seed provides greater flexibility through overdispersion.

However, the adjusted composite-likelihood criteria give a more cautious picture. On the common subset of days where CLAIC and CLBIC could be computed for all six

IVT specifications, the adjusted criteria often favour the simpler Poisson–Exp model. This suggests that the raw likelihood gains from the richer specifications are not always large enough to compensate for their estimated effective complexity. The result is not contradictory. Rather, it highlights that raw in-sample fit and penalized model selection answer different questions.

The adjusted criteria should also be interpreted carefully because the Godambe correction did not complete uniformly across all day-model combinations. The missing values are concentrated in the more numerically demanding specifications, especially NB–Gamma. For this reason, CLAIC and CLBIC were treated as a robustness check rather than as the primary in-sample ranking. Still, the adjusted results are useful because they prevent an overly simple conclusion that the richest model is always preferable. They show that the additional flexibility of the negative-binomial and Gamma specifications comes with an effective complexity cost.

This supports a broader conclusion: the IVT framework is flexible, but this flexibility should be used with care. Richer specifications can improve fit, but they may also be more difficult to estimate, less stable, and less strongly supported after complexity adjustment.

4.3 Forecasting performance

The forecasting results show that in-sample fit does not translate mechanically into forecast dominance. Forecast performance is strongly horizon-dependent.

At the 5-second horizon, the Poisson–INGARCH benchmark performs best for most assets. This suggests that at the shortest horizon, very local observation-driven dynamics are difficult to beat. A simple discrete-time count model can exploit the most recent spread observation effectively, and the additional structure of the IVT models does not automatically give an advantage at this horizon.

At the 1-minute horizon, the IVT models become more competitive. The negative-binomial IVT specifications, especially NB–Gamma and NB–Exp depending on the asset, provide the strongest and most consistent gains relative to the Naive benchmark. This is the horizon where the IVT framework appears most useful for forecasting. The result is intuitive: at one minute, Naive becomes less adequate, but there is still enough dependence for the trawl-based conditional mean to contain useful predictive information.

At the 5-minute and 10-minute horizons, the Naive benchmark becomes difficult to beat. The best IVT competitors are often Gamma-trawl specifications, but the median MSE ratios are generally close to or above one for several assets. This indicates that the predictive value of the fitted trawl dependence structure weakens as the horizon increases. The model-based forecasts mean-revert towards the fitted marginal mean, and this mean reversion does not always improve squared-error performance relative to simple Naive.

Overall, the forecasting results suggest that IVT models are most useful at short and intermediate horizons rather than uniformly across all horizons. The models capture dependence that is relevant for prediction, but the forecasting advantage is not robust enough to dominate simple benchmarks at all horizons. This is important for the practical interpretation of the thesis: the IVT framework is not simply a replacement for benchmark models, but a structured model class that can improve forecasts in specific regimes.

4.4 Numerical and finite-sample considerations

The empirical application also reveals practical challenges in estimating richer IVT specifications. The negative-binomial models are not available on every asset-day because the

moment-based initializer requires the empirical variance to exceed the empirical mean. This issue does not affect the Poisson specifications, but it leads to missing negative-binomial fits for some days of A, DFS, and WM. In addition, the Godambe correction required for CLAIC and CLBIC is numerically demanding, especially for NB–Gamma.

These issues are not merely computational inconveniences. They reflect the fact that richer IVT models can be harder to identify and estimate in finite samples. The simulation study supports this interpretation. Parameter recovery improves with sample size, but the richer specifications, particularly those involving negative-binomial seeds and Gamma trawls, can produce occasional extreme estimates. Mean-based measures such as bias and RMSE can therefore be strongly affected by outlying replications, while median-based errors often show more regular recovery.

This motivates the cautious empirical interpretation used in the thesis. Individual parameter estimates should not be over-interpreted, especially for the richer specifications. More weight should be placed on overall fit, forecast performance, success rates, and robustness across assets and days. From a practical point of view, the results also show that IVT estimation requires safeguards: admissible starting values, checks on optimizer convergence, forecast sanity checks, and careful handling of failed fits.

4.5 Stationarity and intraday variation

Another limitation is the stationarity assumption. Each asset-day is treated as approximately stationary after removing the first hour of trading. This is a reasonable simplification, but the representative-day plot shows that the spread level and volatility can still change during the day. The spread is often higher and more volatile earlier in the sample and lower towards the end of the day.

This remaining intraday variation may help explain why the fitted ACFs do not fully reproduce the empirical dependence structure. A stationary one-factor trawl model has a single marginal distribution and a single dependence profile over the estimation window. If the data contain residual intraday seasonality or local non-stationarity, the fitted model must average over these features. This can lead to imperfect marginal fit, underestimated medium-lag persistence, and horizon-dependent forecasting performance.

One possible extension would be to include deterministic intraday seasonality, either by modelling a time-varying mean or by estimating the IVT model on shorter locally stationary windows. Another possibility would be to use more flexible trawl specifications or multi-factor trawl models that can capture both short-run decay and slower persistence. Such extensions may improve the fit to the empirical ACF, but they would also increase the complexity and numerical burden of the estimation problem.

4.6 Implications and possible extensions

The empirical results suggest several directions for further work. A first extension concerns the specification of the trawl function. In this thesis, the dependence structure is restricted to three parametric families: exponential, IG, and Gamma. This keeps the empirical comparison tractable and interpretable, but it also imposes a specific shape on the autocorrelation decay. The in-sample diagnostics show that the Gamma trawl often improves the ACF fit relative to the exponential trawl, but none of the parametric specifications fully reproduces the empirical medium-lag persistence. This suggests that the remaining dependence misspecification may not only be a matter of choosing between the three parametric families considered here.

A natural way to investigate this limitation would be to use nonparametric trawl estimation as a diagnostic tool. As emphasized by Sauri and Veraart (2022), structural features of the trawl function may be difficult to detect from the autocorrelation function alone, since the ACF is determined by integrated properties of the trawl set. Different trawl functions can therefore imply similar ACFs, while still differing in local shape. In the present application, a nonparametric estimate of the trawl function could be used to assess where the exponential, IG, or Gamma specifications are locally misspecified. This would provide a more direct diagnostic of the dependence structure than comparing fitted parametric ACFs alone.

Such an approach could also motivate hybrid specifications. The bid–ask spread appears to contain both strong short-run dependence and slower medium-lag persistence. A single parametric trawl family may be too restrictive to capture both features simultaneously. One possible extension would therefore be to estimate the trawl function nonparametrically at short lags and then approximate the tail by a simpler parametric form. This would preserve some of the interpretability and forecasting convenience of the parametric IVT framework, while allowing the dependence structure to be guided more directly by the data.

A second extension concerns the estimation procedure. The empirical analysis shows that the richer IVT specifications are more numerically fragile. Negative-binomial specifications are unavailable on some asset-days because the moment-based initializer is not admissible, and the Godambe correction required for CLAIC and CLBIC is particularly difficult for the more flexible models. More robust initialization procedures, alternative optimization strategies, or regularized estimation could reduce the number of failed fits and improve the stability of the richer specifications.

A third extension would be to evaluate the full predictive distribution rather than only conditional-mean forecasts. The IVT framework implies predictive distributions, not just point forecasts. This makes it possible to use density forecast measures, prediction intervals, or proper scoring rules such as the logarithmic score. Such an evaluation would be especially relevant for bid–ask spreads, where occasional large spreads and tail behaviour are important.

Finally, the analysis could be extended beyond the univariate setting. Bid–ask spreads across related assets may respond to common liquidity shocks or market-wide conditions. A multivariate trawl framework could potentially capture such cross-asset dependence. This would be a natural extension of the present thesis, although it would also require substantially more complex estimation and model comparison methods.

5 | Conclusion

This thesis has studied the use of integer-valued trawl processes for modelling and forecasting high-frequency bid–ask spreads. The motivation was that the bid–ask spread is discrete, non-negative, persistent, and often overdispersed. These features make it a natural candidate for an IVT framework, where the Lévy seed determines the marginal distribution and the trawl function determines the serial dependence structure.

The theoretical part of the thesis introduced the relevant trawl-process framework, with particular focus on stationary integer-valued trawl processes, pairwise composite-likelihood estimation, and conditional-mean forecasting. Since the full likelihood is generally not tractable for the models considered, the empirical implementation relied on pairwise composite likelihood. The application compared six IVT specifications, combining Poisson and negative-binomial Lévy seeds with exponential, IG, and Gamma trawl functions. Forecast performance was evaluated against Naive and Poisson–INGARCH benchmarks using an expanding-window design.

The empirical analysis showed that the IVT framework is well suited for describing several important features of bid–ask spread dynamics. In the in-sample analysis, the negative-binomial seed generally improved the marginal fit relative to the Poisson seed, reflecting the importance of overdispersion in the spread distribution. The dependence diagnostics showed that the exponential trawl often decays too quickly, while the IG and especially Gamma trawls provide a better match to the empirical autocorrelation structure. Thus, the empirical results support the interpretation that the seed and trawl components play distinct roles: the seed mainly controls the unconditional distribution, while the trawl function controls the persistence profile.

At the same time, the in-sample results also showed that model ranking depends on the criterion used. The unadjusted composite likelihood often favoured negative-binomial specifications, indicating that the richer models provide stronger raw pairwise fit. However, the adjusted composite-likelihood criteria, computed on the subset of days where the Godambe correction was available for all models, often favoured the simpler Poisson–Exp specification. This suggests that the raw likelihood gains from the richer specifications are not always large enough to compensate for their estimated effective complexity. The empirical evidence therefore does not support a single universally best IVT specification. Instead, it shows that different diagnostics emphasize different aspects of model performance.

The forecasting results were similarly horizon-dependent. At the shortest 5-second horizon, the Poisson–INGARCH benchmark performed best for most assets, showing that simple observation-driven count dynamics are highly competitive for ultra-short-term prediction. At the 1-minute horizon, the negative-binomial IVT specifications performed most consistently, with NB–Gamma and NB–Exp often producing the strongest gains relative to Naive. At the 5-minute and 10-minute horizons, the Naive benchmark became difficult to beat, although Gamma-trawl IVT models remained important competitors. Overall, the IVT forecasts contained useful predictive information, but mainly at short and intermediate horizons rather than uniformly across all forecast horizons.

The thesis also documented important practical limitations. The richer IVT specifications are more numerically fragile than the simpler ones. Negative-binomial fits are unavailable on some asset-days because the moment-based initializer is not admissible, and the Godambe correction required for CLAIC and CLBIC is particularly demanding for some specifications. The simulation study supports the same caution: parameter recovery improves with sample size, but richer negative-binomial and Gamma specifications can

produce occasional extreme estimates in finite samples. For this reason, the empirical interpretation placed more weight on overall fit, forecast performance, success rates, and robustness across days than on individual parameter estimates alone.

A further limitation is the stationarity assumption. Even after removing the first hour of trading, the bid–ask spread can exhibit intraday changes in level and volatility. Since the fitted IVT models are stationary within each estimation window, residual intraday non-stationarity may contribute to the imperfect ACF fit and to the weakening of forecast performance at longer horizons. This points to several possible extensions, including intraday seasonality adjustments, locally stationary specifications, more robust estimation procedures, and nonparametric trawl diagnostics. In particular, nonparametric trawl estimation could be used to assess whether the exponential, IG, and Gamma families are locally misspecified and to guide the construction of more flexible or hybrid trawl specifications.

The main conclusion is therefore that IVT processes are useful for bid–ask spread modelling, but not uniformly dominant as forecasting models. Integer-valued trawl processes provide an interpretable and empirically useful framework for modelling bid–ask spreads. They respect the discreteness of the data, allow flexible marginal distributions, and connect persistence directly to the trawl function. However, their forecasting advantage is horizon-specific, and richer specifications come with increased numerical and effective model complexity. The IVT framework should therefore be viewed not as a uniformly dominant forecasting tool, but as a structured modelling approach that helps separate and analyse the marginal and dependence features of high-frequency spread dynamics.

References

Amihud, Y. and Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics*, 17(2):223–249.

Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. E. D. (2018). *Ambit Stochastics*, volume 88 of *Probability Theory and Stochastic Modelling*. Springer, Cham.

Barndorff-Nielsen, O. E., Lunde, A., Shephard, N., and Veraart, A. E. D. (2014). Integer-valued trawl processes: A class of stationary infinitely divisible processes. *Scandinavian Journal of Statistics*, 41(3):693–724.

Benjamini, Y. and Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1):289–300.

Bennedsen, M., Lunde, A., Shephard, N., and Veraart, A. E. D. (2023). Inference and forecasting for continuous-time integer-valued trawl processes. *Journal of Econometrics*, 236:105476.

Diebold, F. X. and Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3):253–263.

Huang, R. D. and Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies*, 10(4):995–1034.

Liboschik, T., Fokianos, K., and Fried, R. (2017). tscount: An R package for analysis of count time series following generalized linear models. *Journal of Statistical Software*, 82(5):1–51.

Madhavan, A. (2000). Market microstructure: A survey. *Journal of Financial Markets*, 3(3):205–258.

Rajput, B. S. and Rosiński, J. (1989). Spectral representations of infinitely divisible processes. *Probability Theory and Related Fields*, 82(3):451–487.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of Finance*, 39(4):1127–1139.

Sauri, O. and Veraart, A. E. D. (2022). Nonparametric estimation of trawl processes: Theory and applications.

Stoll, H. R. (1989). Inferring the components of the bid-ask spread: Theory and empirical tests. *The Journal of Finance*, 44(1):115–134.

A | Appendix

A.1 Results for the simulation study

This appendix reports the parameter-recovery plots from the simulation study discussed in Chapter 3. The study is a local finite-sample exercise based on empirically calibrated parameter vectors, and is therefore intended as an application-specific check of the estimation procedure rather than a global analysis over the full parameter space.

The design is asset-specific. For each of the four assets A, WAT, DFS, and WM, we consider the six IVT specifications used in the empirical analysis: Poisson–Exp, Poisson–IG, Poisson–Gamma, NB–Exp, NB–IG, and NB–Gamma. For each asset–model combination, the data-generating parameters are obtained as componentwise medians of day-level GMM-style calibration vectors computed from the empirical training windows. For each scenario, 1000 independent replications are simulated at sample sizes

$$n \in \{1000,2500,5000,7500\},$$

and each simulated sample is re-estimated using the same composite-likelihood procedure as in the empirical analysis. The plots report parameter-specific mean bias and RMSE across successful replications, with the Poisson seed parameter denoted by μ as in the main text.

The simulation runs are generally stable. All Poisson specifications are estimated successfully in all replications, and most negative-binomial specifications also have full success rates. The few numerical failures are concentrated in NB–Gamma, with the lowest success rate being 98.9% for asset A at $n = 1000$. The Poisson-based models show comparatively regular recovery, with moderate mean bias and RMSE that generally decrease as the sample size increases. The negative-binomial specifications are more delicate: rare extreme estimates, especially of the size parameter m, can dominate mean-based measures even when the central tendency is more stable.

The plots should therefore be interpreted with this asymmetry in mind. For Poisson-based models, the mean bias and RMSE curves give a direct indication of parameter recovery. For negative-binomial models, they also reflect occasional finite-sample identification problems in the parametrisation, where m, p, and the trawl scale parameters can partially compensate each other. This supports the interpretation used in the empirical analysis: richer IVT specifications can improve fit and forecasting performance, but they are also more numerically delicate, and individual parameter estimates should not be over-interpreted.

Table A.1: Asset-specific simulation scenarios. The data-generating parameters are componentwise medians of day-level calibration vectors computed from the empirical training windows. For negative-binomial specifications, the number of source windows may be smaller because the moment-based negative-binomial initializer is only admissible when the empirical variance exceeds the empirical mean.

Asset	Model	Source windows	Data-generating parameters
A	Poisson–Exp	458	$\mu = 14.141$, $\lambda = 2.224$
A	Poisson–IG	458	$\mu = 19.654$, $\delta = 0.912$, $\gamma = 0.303$
A	Poisson–Gamma	458	$\mu = 17.815$, $H = 0.433$, $\alpha = 0.157$
A	NB–Exp	433	$m = 27.626$, $p = 0.366$, $\lambda = 2.131$
A	NB–IG	433	$m = 38.873$, $p = 0.366$, $\delta = 0.882$, $\gamma = 0.307$
A	NB–Gamma	433	$m = 34.979$, $p = 0.366$, $H = 0.430$, $\alpha = 0.160$
WAT	Poisson–Exp	415	$\mu = 31.563$, $\lambda = 0.891$
WAT	Poisson–IG	415	$\mu = 36.870$, $\delta = 0.458$, $\gamma = 0.436$
WAT	Poisson–Gamma	415	$\mu = 35.993$, $H = 0.265$, $\alpha = 0.261$
WAT	NB–Exp	415	$m = 8.504$, $p = 0.797$, $\lambda = 0.891$
WAT	NB–IG	415	$m = 10.101$, $p = 0.797$, $\delta = 0.458$, $\gamma = 0.436$
WAT	NB–Gamma	415	$m = 9.736$, $p = 0.797$, $H = 0.265$, $\alpha = 0.261$
DFS	Poisson–Exp	452	$\mu = 14.640$, $\lambda = 3.487$
DFS	Poisson–IG	452	$\mu = 26.887$, $\delta = 1.101$, $\gamma = 0.179$
DFS	Poisson–Gamma	452	$\mu = 21.314$, $H = 0.438$, $\alpha = 0.088$
DFS	NB–Exp	390	$m = 30.177$, $p = 0.382$, $\lambda = 3.337$
DFS	NB–IG	390	$m = 67.599$, $p = 0.382$, $\delta = 1.027$, $\gamma = 0.179$
DFS	NB–Gamma	390	$m = 46.500$, $p = 0.382$, $H = 0.426$, $\alpha = 0.088$
WM	Poisson–Exp	437	$\mu = 15.114$, $\lambda = 2.934$
WM	Poisson–IG	437	$\mu = 21.847$, $\delta = 1.145$, $\gamma = 0.264$
WM	Poisson–Gamma	437	$\mu = 19.472$, $H = 0.530$, $\alpha = 0.132$
WM	NB–Exp	391	$m = 30.585$, $p = 0.363$, $\lambda = 2.908$
WM	NB–IG	391	$m = 52.308$, $p = 0.363$, $\delta = 1.145$, $\gamma = 0.270$
WM	NB–Gamma	391	$m = 41.100$, $p = 0.363$, $H = 0.529$, $\alpha = 0.136$

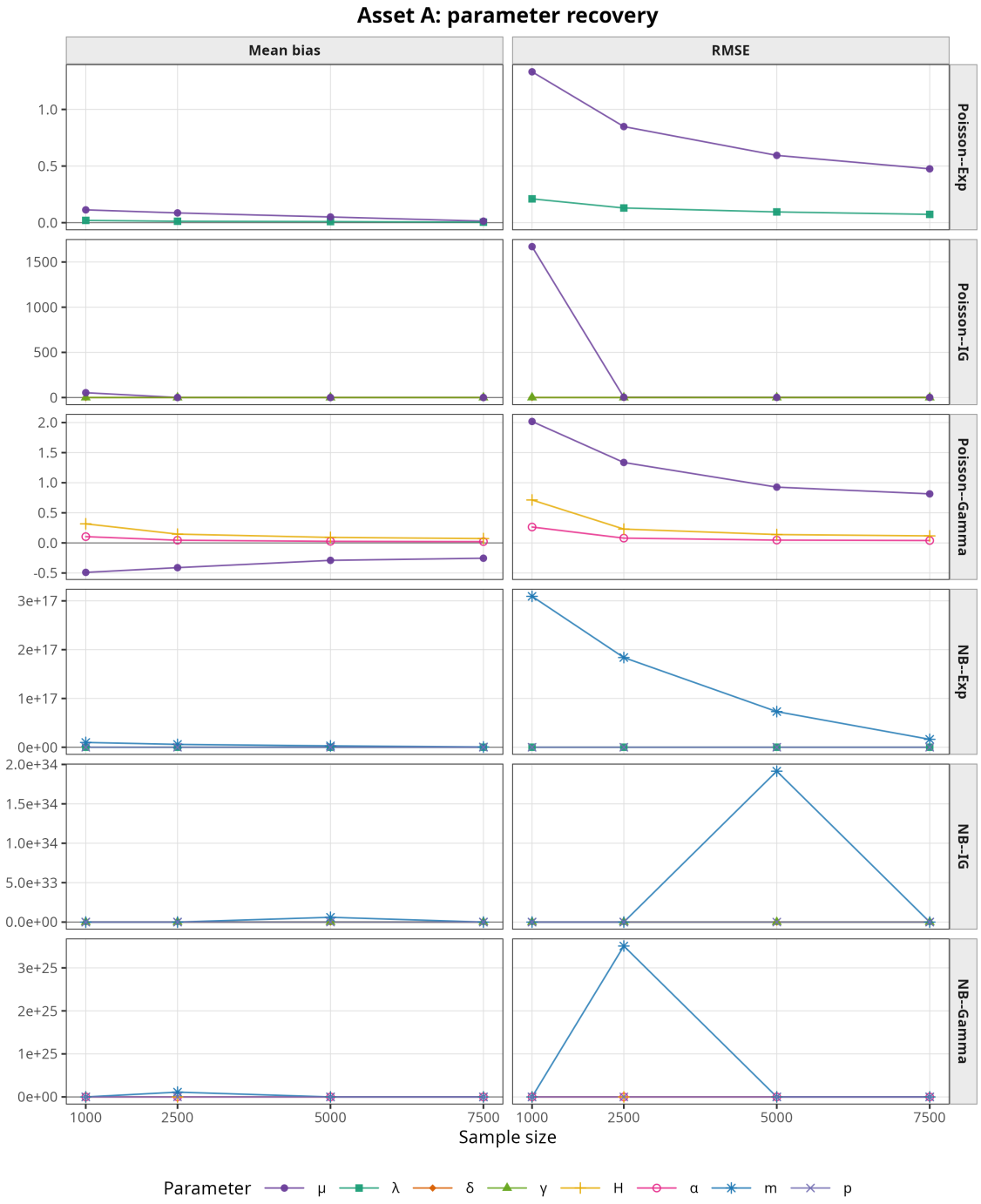


Figure A.2: Parameter-recovery results for asset A. Each panel reports mean bias and RMSE across sample sizes for one empirically calibrated IVT scenario, based on 1000 Monte Carlo replications.

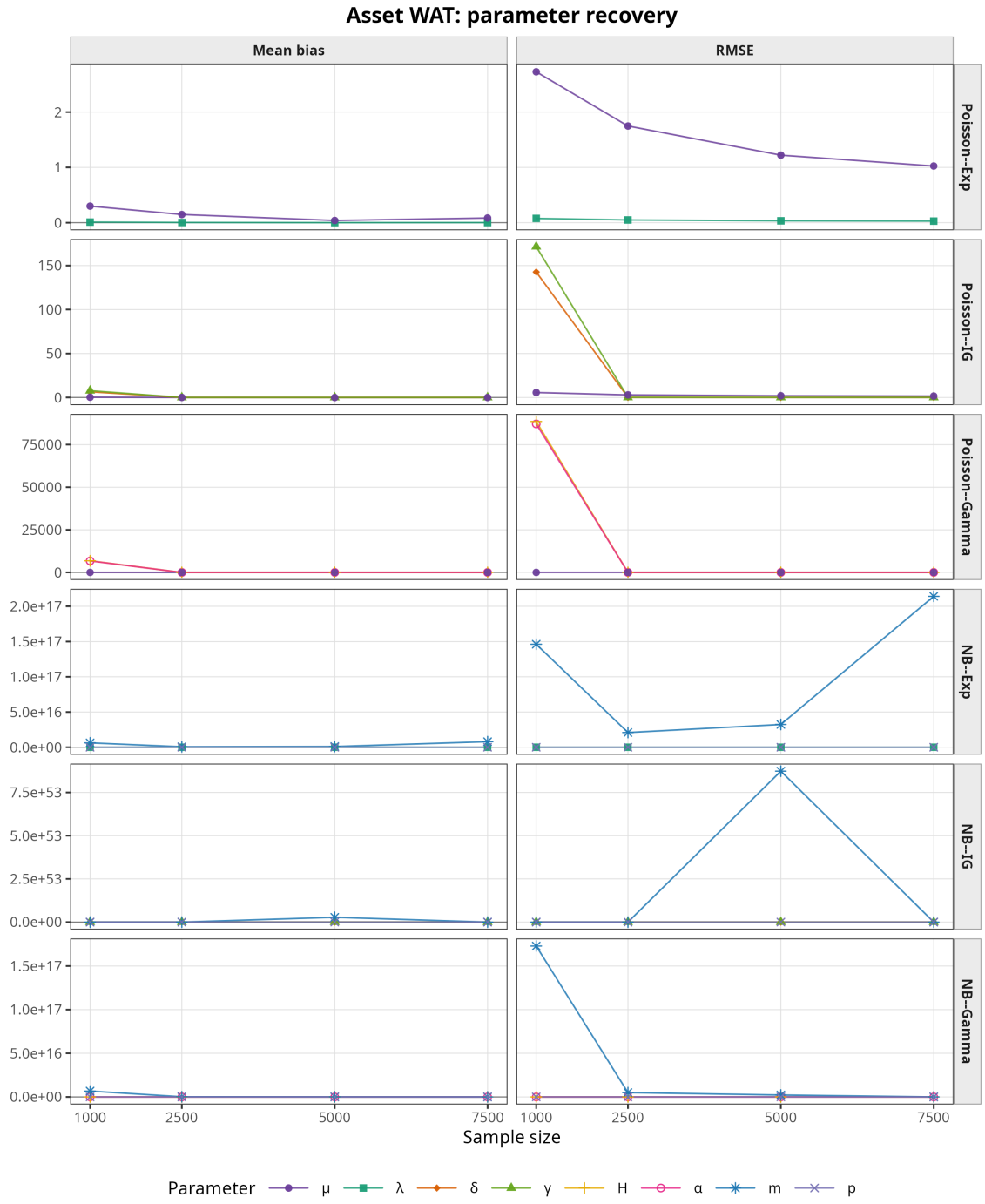


Figure A.3: Parameter-recovery results for asset WAT. Each panel reports mean bias and RMSE across sample sizes for one empirically calibrated IVT scenario, based on 1000 Monte Carlo replications.

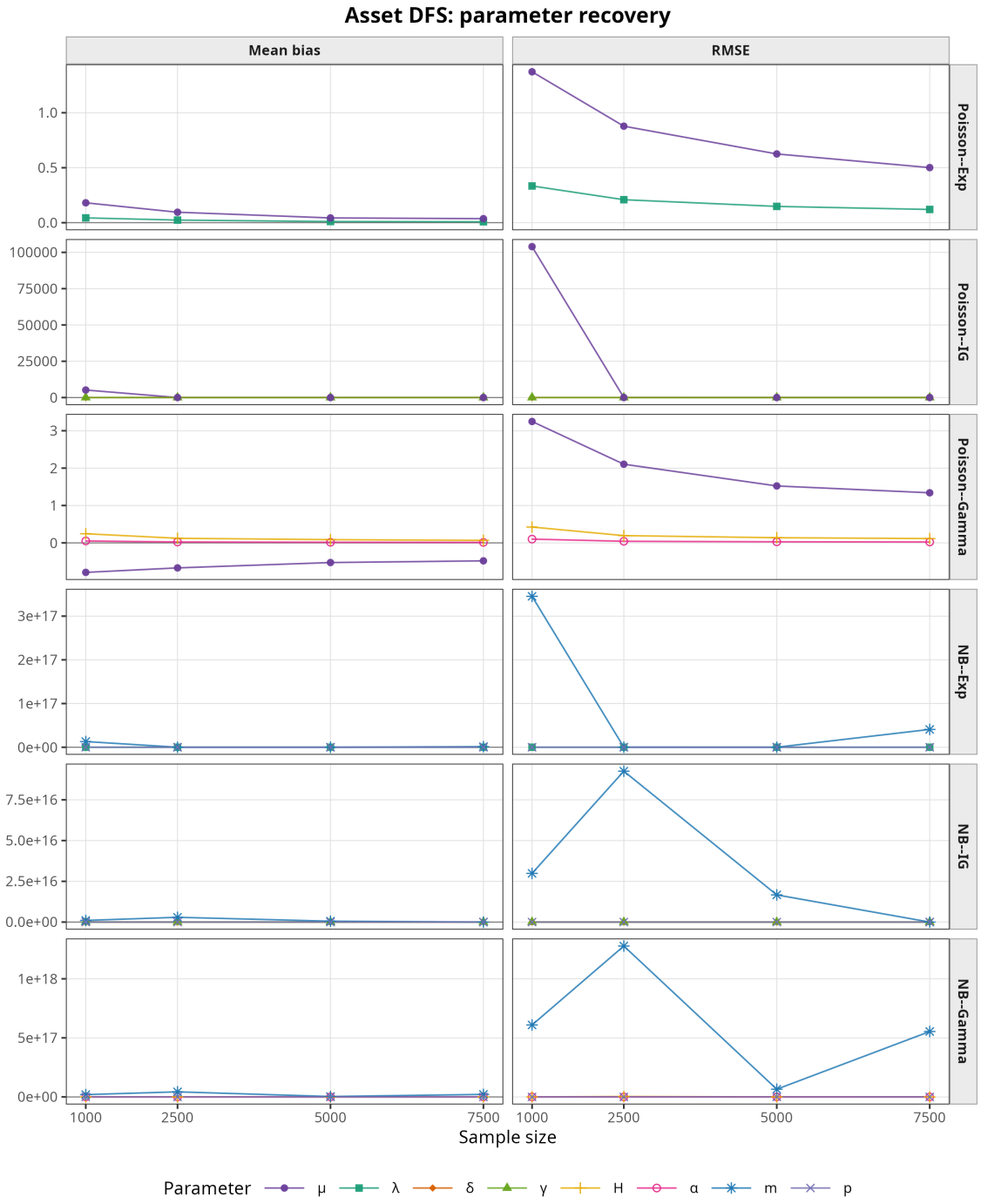


Figure A.4: Parameter-recovery results for asset DFS. Each panel reports mean bias and RMSE across sample sizes for one empirically calibrated IVT scenario, based on 1000 Monte Carlo replications.

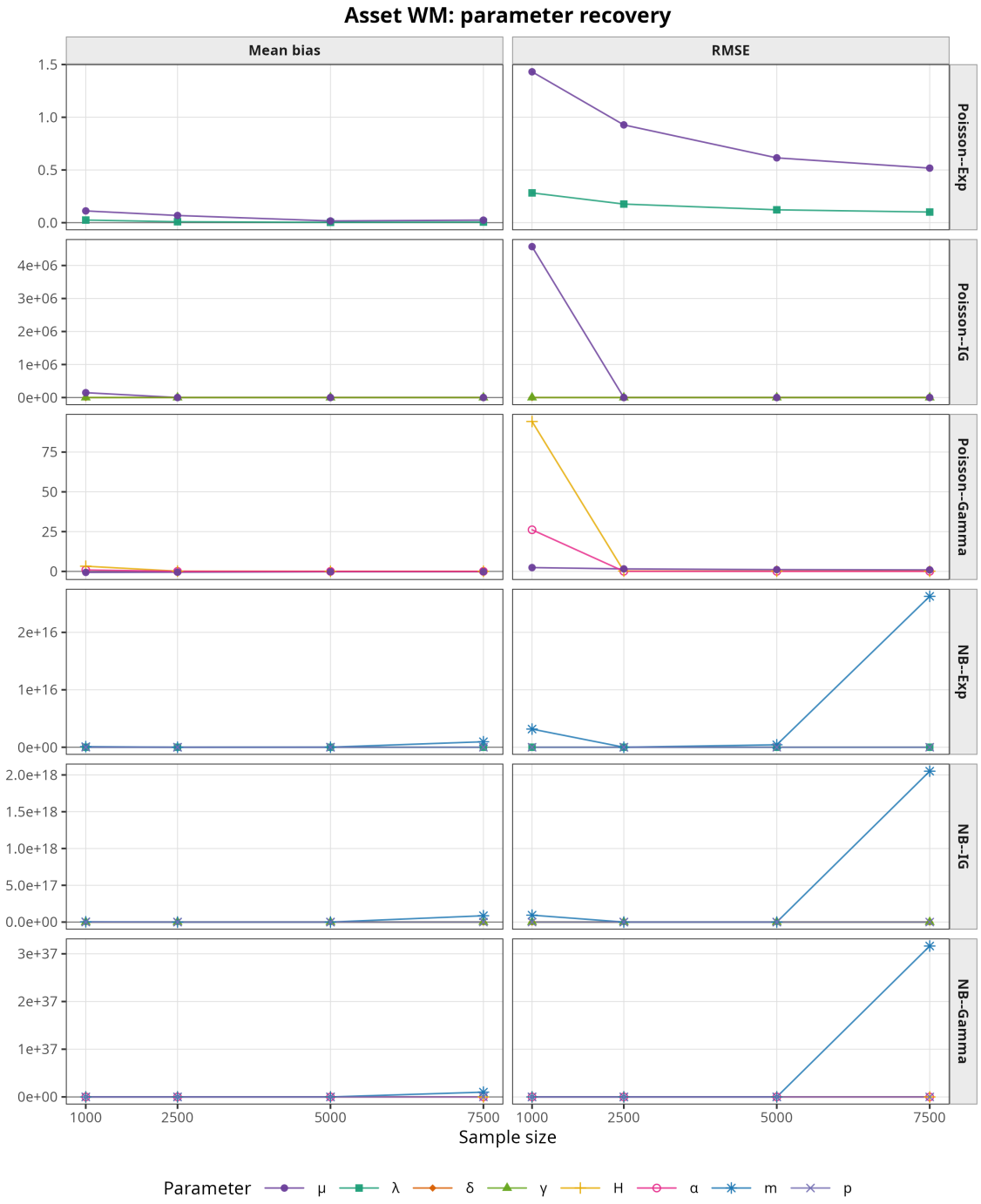


Figure A.5: Parameter-recovery results for asset WM. Each panel reports mean bias and RMSE across sample sizes for one empirically calibrated IVT scenario, based on 1000 Monte Carlo replications.

A.2 Multi-day CLAIC and CLBIC calculations

For the multi-day in-sample comparison, we attempted to compute the adjusted composite-likelihood criteria CLAIC and CLBIC for all fitted day-model combinations. The calculation requires the Godambe correction and is substantially more expensive than the unadjusted composite likelihood. In practice, the calculation completed for most, but not all, fitted combinations. Table A.6 reports both the first-pass fit success rate and the completion rate for the adjusted criteria.

The missing adjusted criteria are not distributed uniformly across specifications. The most difficult cases are concentrated in the richer negative-binomial specifications, especially NB–Gamma. For this reason, the adjusted criteria are not used as the primary multi-day ranking in Section 3.2.2. Instead, they are reported here as a robustness check.

To obtain a comparable adjusted ranking, Table A.7 reports CLAIC and CLBIC winner frequencies only on the subset of asset-days for which all six specifications have finite CLAIC and CLBIC values. This gives a common candidate set within each asset, but it also reduces the number of usable days. The adjusted rankings should therefore be interpreted as a robustness check rather than as the main in-sample comparison.

Table A.6: Completion of multi-day CLAIC/CLBIC calculations by asset and model. Fit success reports the share of asset-days for which the first-pass IVT estimation succeeded. Adjusted completion reports the share of asset-days for which both CLAIC and CLBIC were recovered.

Asset	Model	Total days	Fit success	Adjusted completed	Adjusted completion
A	Poisson–Exp	458	458	458	100.0%
A	Poisson–IG	458	458	445	97.2%
A	Poisson–Gamma	458	458	439	95.9%
A	NB–Exp	458	433	433	94.5%
A	NB–IG	458	433	411	89.7%
A	NB–Gamma	458	433	366	79.9%
WAT	Poisson–Exp	415	415	415	100.0%
WAT	Poisson–IG	415	415	402	96.9%
WAT	Poisson–Gamma	415	415	415	100.0%
WAT	NB–Exp	415	415	415	100.0%
WAT	NB–IG	415	415	376	90.6%
WAT	NB–Gamma	415	415	280	67.5%
DFS	Poisson–Exp	452	452	452	100.0%
DFS	Poisson–IG	452	452	378	83.6%
DFS	Poisson–Gamma	452	452	439	97.1%
DFS	NB–Exp	452	390	390	86.3%
DFS	NB–IG	452	390	382	84.5%
DFS	NB–Gamma	452	390	363	80.3%
WM	Poisson–Exp	437	437	437	100.0%
WM	Poisson–IG	437	437	394	90.2%
WM	Poisson–Gamma	437	437	417	95.4%
WM	NB–Exp	437	391	391	89.5%
WM	NB–IG	437	391	374	85.6%
WM	NB–Gamma	437	391	341	78.0%

The first-pass fit success rates show that the Poisson specifications are estimated on all available days, whereas the negative-binomial specifications fail on some days for A,

DFS, and WM. These failures occur when the moment-based initialization for the negative-binomial seed is not admissible, typically because the empirical variance is not sufficiently larger than the empirical mean. Conditional on first-pass fit success, the adjusted-criterion calculation is still most difficult for the more flexible specifications. This is particularly visible for NB–Gamma, where the adjusted completion rate is lower than for the other models.

Table A.7: Adjusted composite-likelihood rankings on the common subset. The common subset contains only asset-days for which all six IVT specifications have finite CLAIC and CLBIC values. Percentages report winner frequencies within this reduced subset.

Asset	Criterion	Common days	Most frequent winner	Second most frequent winner
A	CLAIC	332	Poisson–Exp (77.4%)	NB–IG (20.5%)
A	CLBIC	332	Poisson–Exp (74.1%)	NB–IG (23.2%)
WAT	CLAIC	255	Poisson–Exp (80.4%)	NB–IG (18.4%)
WAT	CLBIC	255	Poisson–Exp (75.7%)	NB–IG (23.1%)
DFS	CLAIC	286	Poisson–Exp (84.6%)	NB–IG (10.8%)
DFS	CLBIC	286	Poisson–Exp (83.9%)	NB–IG (11.5%)
WM	CLAIC	281	Poisson–Exp (79.0%)	NB–IG (17.8%)
WM	CLBIC	281	Poisson–Exp (77.9%)	NB–IG (18.5%)

The common-subset adjusted rankings differ substantially from the unadjusted composite-likelihood rankings in Section 3.2.2. Whereas the unadjusted composite likelihood tends to favour negative-binomial specifications, the adjusted criteria most often select the simpler Poisson–Exp model. This suggests that the raw likelihood gains from the richer specifications are not always large enough to compensate for their estimated effective complexity.

At the same time, the adjusted ranking should not be interpreted in isolation. It is based on a reduced subset of days, and the reduction is partly caused by the numerical difficulty of computing the Godambe correction for the more flexible specifications. The adjusted criteria therefore reinforce the need to consider several diagnostics jointly: unadjusted composite likelihood for raw joint fit, marginal RMSE for the unconditional distribution, ACF RMSE for dependence fit, and out-of-sample forecast performance for predictive usefulness.

A.3 Multi-day Diebold–Mariano heatmaps

This appendix reports the full multi-day pairwise Diebold–Mariano heatmaps used in the forecasting analysis. The tests are computed separately for each trading day, forecast horizon, asset, and pair of models, using squared forecast errors as the loss function. The heatmaps summarize these day-level tests by recording which model obtains significantly lower squared-error loss more often across trading days.

In each off-diagonal cell, the label gives the model with the larger share of significant day-level wins in the pairwise comparison, and the percentage gives the share of valid trading days on which that model has significantly lower loss. Percentages below 50% may occur because many days can be non-significant, or because significant wins are split between the two models. Blue cells indicate that the row model has the larger share of significant wins, while red cells indicate that the column model has the larger share. Grey cells correspond to the diagonal.

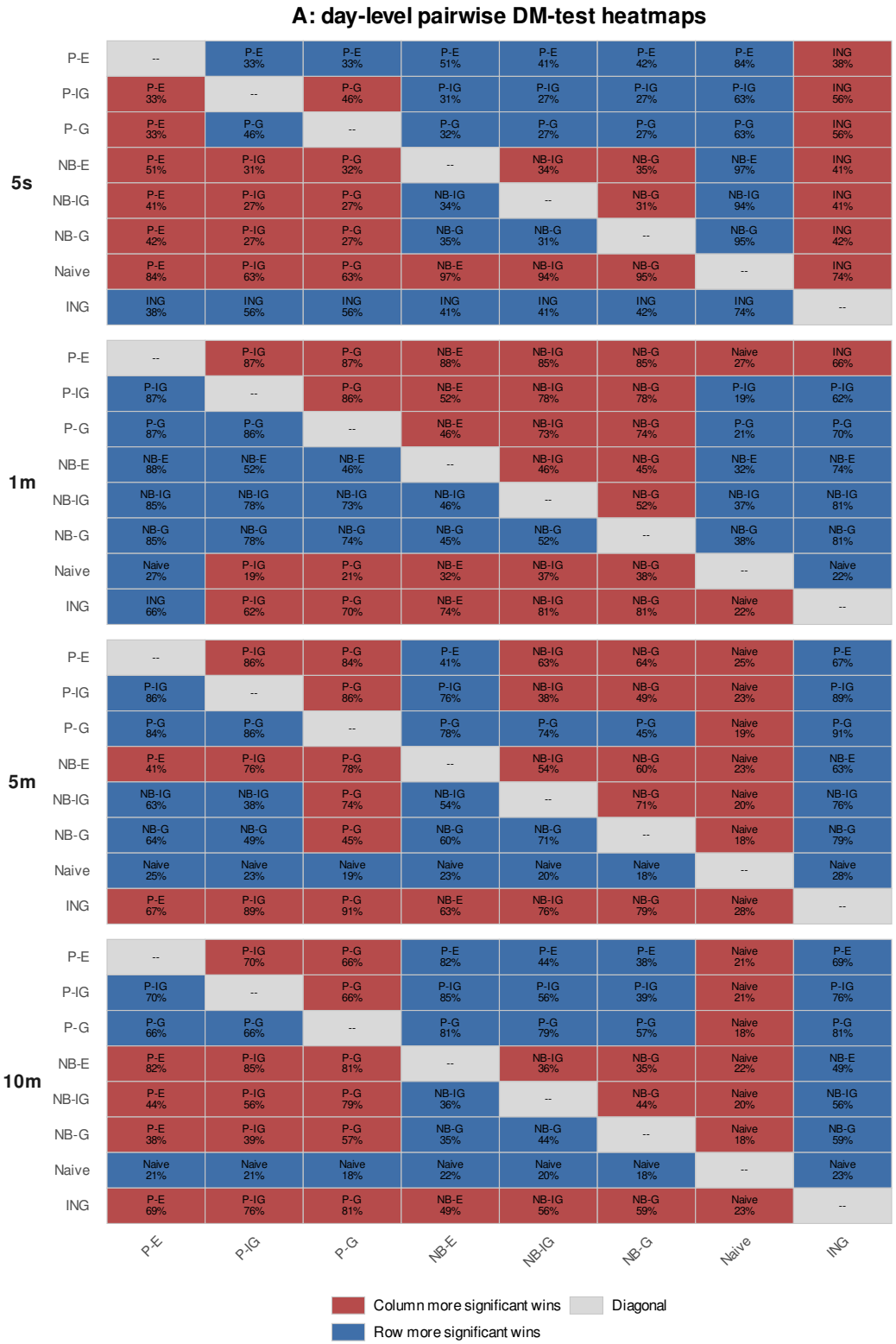


Figure A.8: Multi-day pairwise Diebold–Mariano heatmap for asset A. Each cell reports the model with the larger share of significant day-level wins in the pairwise comparison. The percentage is the share of valid trading days on which the labelled model has significantly lower squared-error loss.

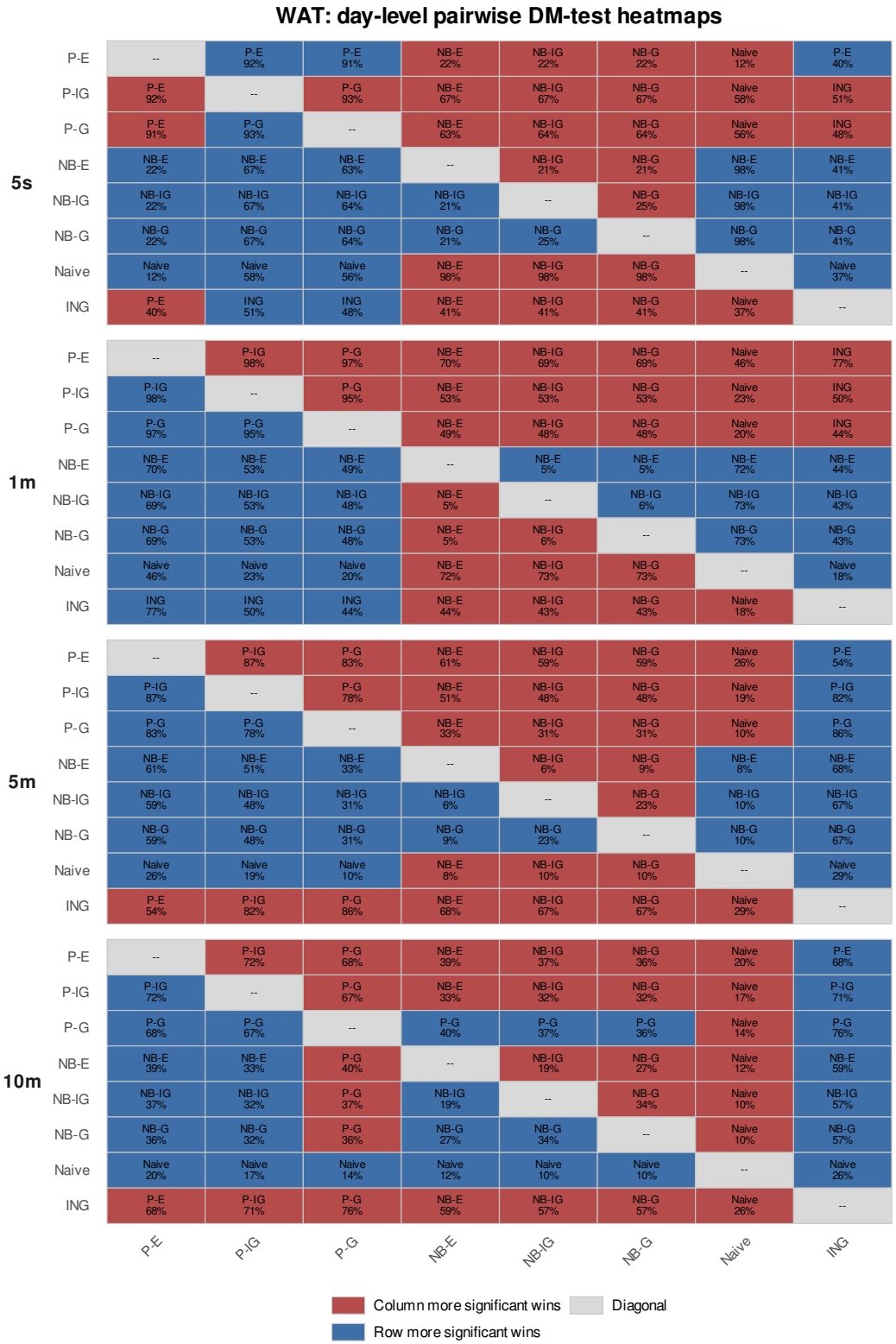


Figure A.9: Multi-day pairwise Diebold–Mariano heatmap for asset WAT. Each cell reports the model with the larger share of significant day-level wins in the pairwise comparison. The percentage is the share of valid trading days on which the labelled model has significantly lower squared-error loss.

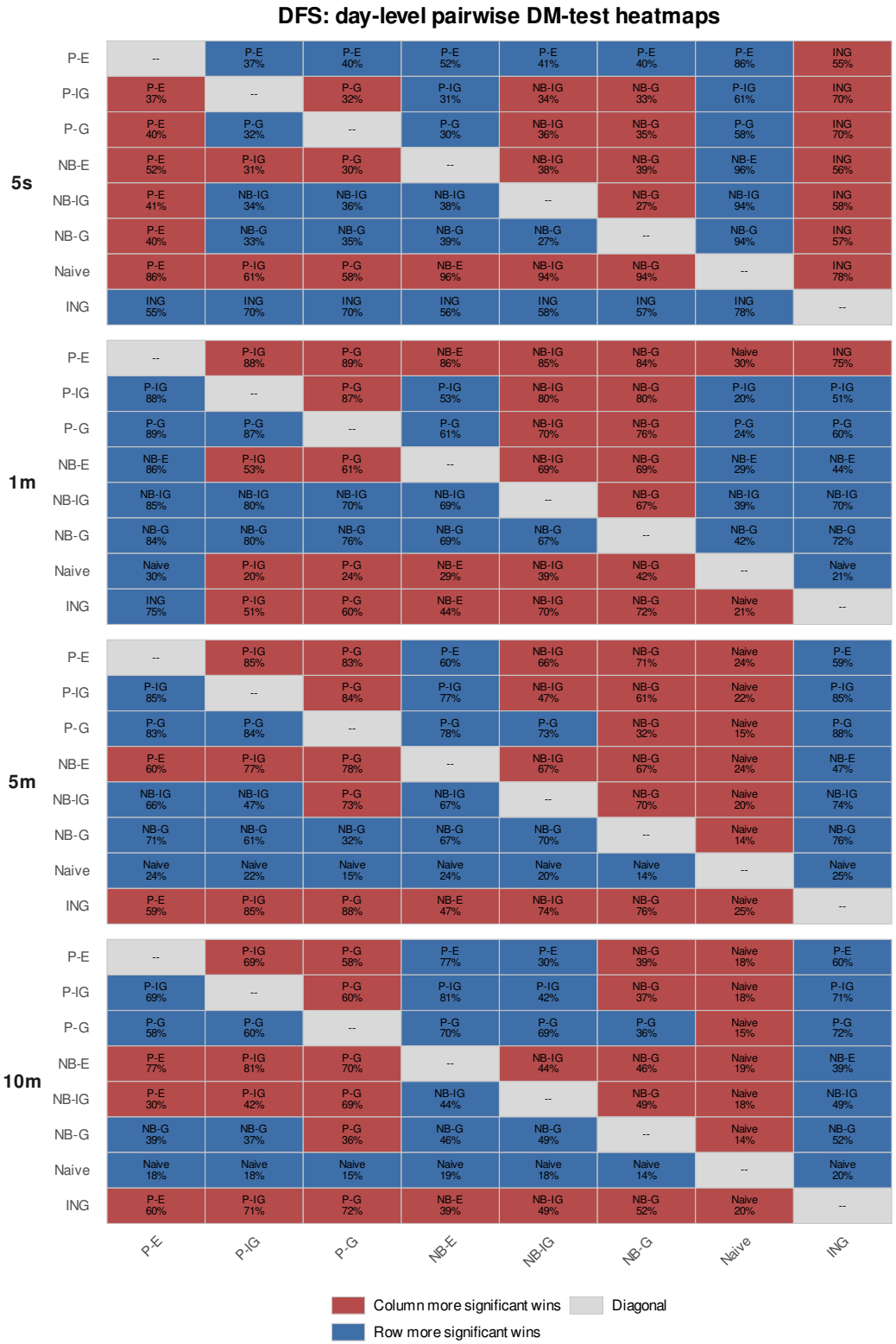


Figure A.10: Multi-day pairwise Diebold–Mariano heatmap for asset DFS. Each cell reports the model with the larger share of significant day-level wins in the pairwise comparison. The percentage is the share of valid trading days on which the labelled model has significantly lower squared-error loss.

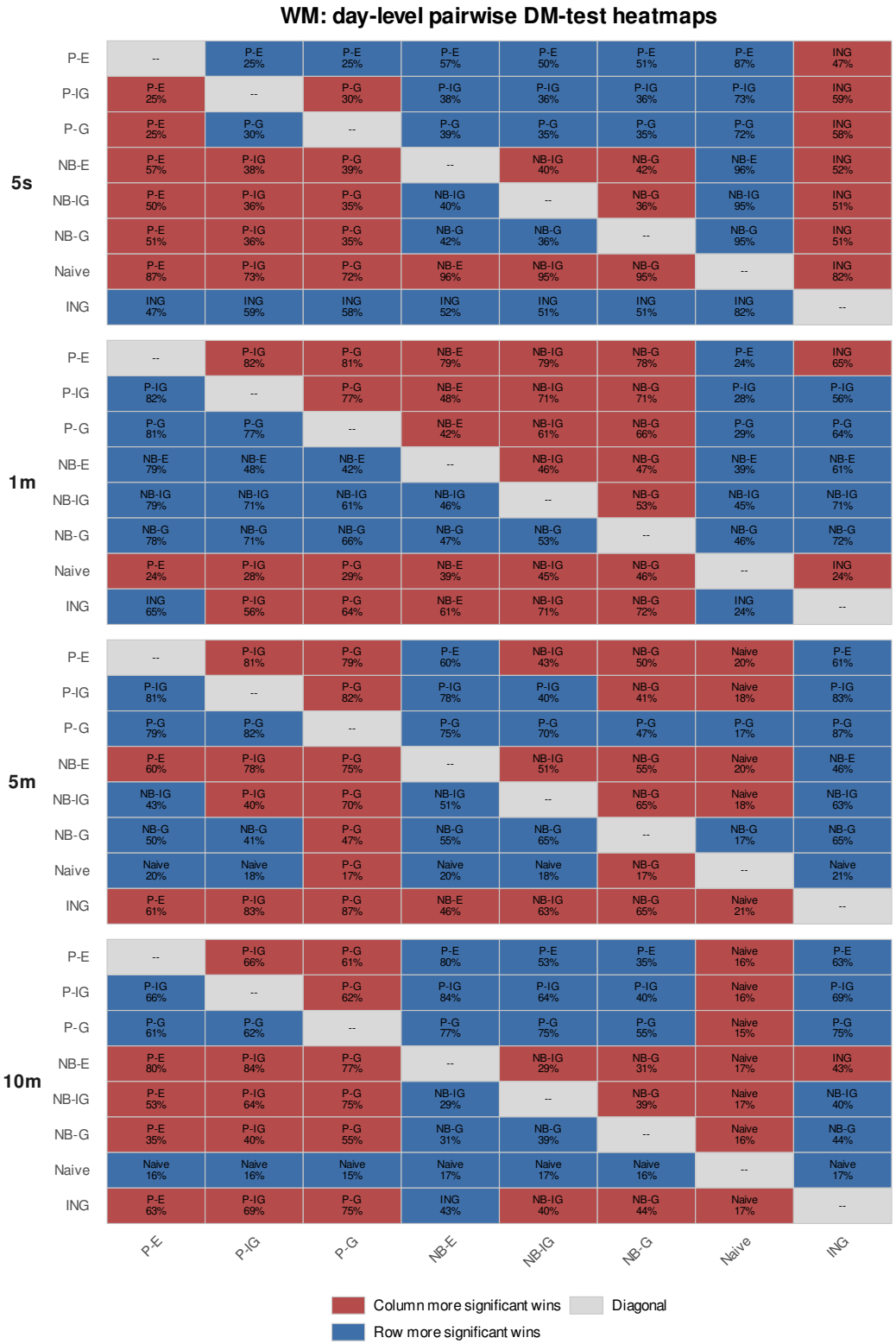


Figure A.11: Multi-day pairwise Diebold–Mariano heatmap for asset WM. Each cell reports the model with the larger share of significant day-level wins in the pairwise comparison. The percentage is the share of valid trading days on which the labelled model has significantly lower squared-error loss.