



## Detecting toxic flow

Álvaro Cartea, Gerardo Duran-Martin & Leandro Sánchez-Betancourt

To cite this article: Álvaro Cartea, Gerardo Duran-Martin & Leandro Sánchez-Betancourt (11 Feb 2026): Detecting toxic flow, Quantitative Finance, DOI: [10.1080/14697688.2026.2619539](https://doi.org/10.1080/14697688.2026.2619539)

To link to this article: <https://doi.org/10.1080/14697688.2026.2619539>



© 2026 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 11 Feb 2026.



Submit your article to this journal [↗](#)



Article views: 58



View related articles [↗](#)



View Crossmark data [↗](#)

# Detecting toxic flow

ÁLVARO CARTEA <sup>†‡</sup>, GERARDO DURAN-MARTIN <sup>‡</sup> and LEANDRO SÁNCHEZ-BETANCOURT <sup>†‡\*</sup>

<sup>†</sup>Mathematical Institute, University of Oxford, Oxford, UK  
<sup>‡</sup>Oxford-Man Institute of Quantitative Finance, Oxford, UK

(Received 9 June 2025; accepted 15 January 2026)

This paper develops a framework to predict toxic trades that a broker receives from her clients. Toxic trades are predicted with a novel online learning Bayesian method which we call the *projection-based unification of last-layer and subspace estimation* (PULSE). PULSE is a fast and statistically-efficient Bayesian procedure for online training of neural networks. We employ a proprietary dataset of foreign exchange transactions to test our methodology. Neural networks trained with PULSE outperform standard machine learning and statistical methods when predicting if a trade will be toxic; the benchmark methods are logistic regression, random forests, and a recursively-updated maximum-likelihood estimator. We devise a strategy for the broker who uses toxicity predictions to internalise or to externalise each trade received from her clients. Our methodology can be implemented in real-time because it takes less than one millisecond to update parameters and make a prediction. Compared with the benchmarks, online learning of a neural network with PULSE attains the highest PnL and avoids the most losses by externalising toxic trades.

**Keywords:** Brokers; Toxic flow; Detection; Kalman filter; Online learning; Algorithmic trading; Bayesian neural networks

## 1. Introduction

Liquidity providers are key to well-functioning financial markets. In foreign exchange (FX), as in other asset classes, broker-client relationships are ubiquitous. The broker streams bid and ask quotes to her clients and the clients decide when to trade on these quotes, so the broker bears the risk of adverse selection when trading with better informed clients. These risks are borne by both liquidity providers who stream quotes to individual parties and by market participants who provide liquidity in the books of electronic exchanges. However, in contrast to electronic order books in which trading is anonymous for all participants (e.g. Nasdaq, LSE, Euronext), in broker-client relationships the broker knows which client executed the order. This privileged information can be used by the broker to classify flow, i.e. toxic or benign, and to devise strategies that mitigate adverse selection costs.

In the literature, models generally classify traders as informed or uninformed; see e.g. Bagehot (1971), Copeland and Galai (1983), Grossman and Stiglitz (1980), Amihud and Mendelson (1980), Kyle (1985, 1989), and Glosten and Milgrom (1985). In equity markets, many studies focus on informed flow (i.e. asymmetry of information) across various

traded stocks, see e.g. Easley *et al.* (1996) who study the probability of informed trading at the stock level, while our study focuses on each trade because we have trader identification. In FX markets, Butz and Oomen (2019) develop a model for internalisation, and Oomen (2017) studies execution in an FX aggregator and the market impact of internalisation-externalisation strategies. Overall, studies of toxic flow and information asymmetry do not make predictions of toxicity at the trade level. To the best of our knowledge, ours is the first paper in the literature to use FX data with trader identification to predict the toxicity of each trade.

In our work, a trade is toxic if a client can unwind the trade within a given time window and make a profit (i.e. a loss for the broker). Toxic trades are not necessarily informed, nor informed trades are necessarily toxic. An uninformed client can execute a trade that becomes toxic for the broker because of the random fluctuations of exchange rates. Ultimately, the broker's objective is to avoid holding loss-leading trades in her books, so it is more effective to focus on market features and on each trade the broker fills, rather than on whether a particular client is classified as informed or uninformed. For simplicity, theoretical models in the literature assume traders are informed or uninformed, while in practice not all trades sent by one particular client are motivated by superior information.

\*Corresponding author. Email: [sanchezbetan@maths.ox.ac.uk](mailto:sanchezbetan@maths.ox.ac.uk)

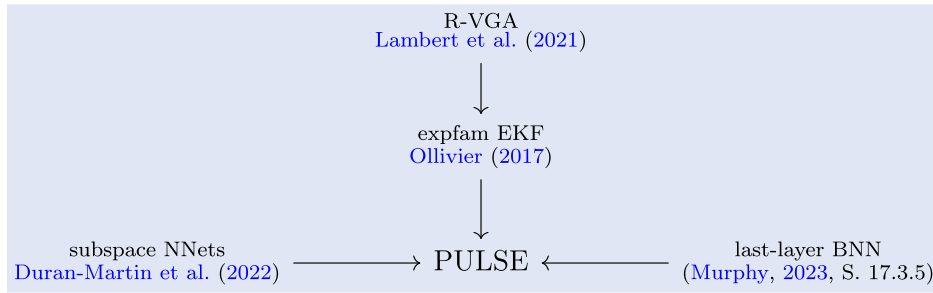


Figure 1. Relationship of PULSE to other models.

The main contributions of our paper are as follows. We predict the toxicity of each incoming trade with machine learning and statistical methods, such as logistic regression, random forests, a recursively updated maximum-likelihood estimator, and a neural network (NNet). We devise a novel algorithm to update the parameters of the NNet sequentially; we call this rule of learning PULSE, which stands for projection-based unification of last-layer and subspace estimation. We deploy our toxicity prediction models in a proprietary dataset, and we find that using a single model for all clients (employing client-specific features) outperforms the use of one model per client. We also find that, compared with the benchmarks, the methodology we put forward attains the highest PnL and avoids the most losses by externalising toxic trades.

Our new method employs a NNet to compute the probability that a trade will be toxic. After the outcome of each trade, toxic or benign, PULSE updates the parameters of the NNet. To update the parameters efficiently at each timestep, PULSE follows three steps: one, split the last layer from the feature-transformation layers of a NNet; two, project the parameters of the feature-transformation layers onto an affine subspace; and three, devise a recursive formula to estimate a posterior distribution over the projected feature-transformation parameters and last-layer parameters. Specifically, we extend the subspace NNet model (subspace NNets) of Duran-Martin *et al.* (2022) to classification tasks. We also use the exponential-family extended Kalman filter (expfam EKF) method of Ollivier (2018) and we follow the ideas of the recursive variational Gaussian approximation (R-VGA) results of Lambert *et al.* (2021) to obtain the update equations in PULSE. Finally, we impose a prior independence between the hidden layers of the NNet and the output layer, extending the work in the last-layer Bayesian NNet (last-layer BNNs). Figure 1 shows the relationship of PULSE to previous methods. In short, PULSE is a statistically-efficient update rule to learn the parameters of a NNet sequentially.

To evaluate the predictive performance of our model and the efficacy of the broker’s strategy, we use a proprietary dataset of FX transactions from 28 June 2022 to 21 October 2022. Initially, the models are trained with data between 28 June and 31 July, and the remainder of the data (1 August to 21 October) is used to deploy the strategy, i.e. use predictions of toxicity for each trade to inform the internalisation-externalisation strategy we develop. During the deploy phase, the maximum-likelihood estimator is continually updated with the running average of the toxic trades and the parameters of the NNet are updated with PULSE, while the models based on logistic regression and random

forests are not updated.† For a given toxicity horizon and a cutoff probability, the strategy internalises the trade if the probability that the trade is toxic is less than or equal to the cutoff probability, otherwise it externalises the trade. We compute the PnL of all trades that the broker internalised and the losses she avoids by externalising trades. We find that a NNet trained with PULSE delivers the best combination of PnL and avoided loss across all toxicity horizons we consider in this paper.

Finally, we find that a universal model is more advantageous than one model per trader. That is, we obtain higher accuracies (when predicting toxic trades) when we train one model for all traders than when we train one model per trader; higher accuracies result from having more data. When one restricts to one model per trader, the model for traders with fewer transactions is outperformed by a universal model that is trained on more datapoints. We also find that if we build a universal model that does not consider the inventory, cash, and recent activity of clients (i.e. the broker does not use the identification of the trader), the performance of PULSE deteriorates substantially when compared to the performance of a model that includes the identity and unique features of the trader. Thus, in our dataset, client-specific variables does add value to predict the toxicity of trades.

The remainder of paper is organised as follows. Section 2 describes the data. Section 3 defines toxicity, provides statistics about the clients in the dataset, and illustrates the toxicity profiles of clients. Section 4 introduces PULSE, which is a fast and statistically-efficient Bayesian procedure for online training of neural networks. Section 5 shows implementation details. Section 6 uses proprietary datasets to evaluate PULSE against alternative methods. Section 7 presents conclusions. We collect proofs, together with additional robustness checks, in the appendix.

## 2. Data and preliminary analysis

We employ data for the currency pair EUR/USD from LMAX Broker and from LMAX Exchange for the period 28 June 2022 to 21 October 2022.‡ For each liquidity taking trade filled by the broker, we use the direction of trade (buy or

† Robustness checks in the appendix study the performance of the logistic regression and random forests when the models are re-trained weekly.

‡ www.lmax.com.

Table 1. Trading activity in the pair EUR/USD between clients and LMAX Broker over the period 28 June 2022 to 21 October 2022. Volumes are reported in one hundred million euros.

	Number of trades	Total volume in €100,000,000	Avg daily volume
Client 1	312,073	43.702	0.520
Client 2	56,705	3.006	0.036
Client 3	28,185	3.278	0.039
Client 4	27,743	0.456	0.005
Client 5	23,938	27.483	0.348
Client 6	13,379	5.379	0.064
Total	462,023	83.304	1.012

sell), the timestamp when LMAX Broker processed the trade, and the volume of the trade. Also, we use the best quotes and volumes available in LMAX Exchange at a microsecond frequency. In contrast to LMAX Broker, traders who interact in the limit order book (LOB) of LMAX Exchange do not know the identity of their counterparties. The LOB uses price-time priority to clear supply and demand of liquidity — as in traditional electronic order books in equity markets, such as those of Nasdaq, Euronext, and the London Stock Exchange.

Table 1 shows summary statistics for the trading activity of six clients of LMAX Broker in the pair EUR/USD.

Below, we work with the data of Clients 1 to 6 and we assume that the broker quotes her clients the best available rates in LMAX Exchange net of fees. Transaction costs in FX are around \$3 per million euros traded (see e.g. Cartea and Sánchez-Betancourt 2023), so this assumption provides clients with a discount of \$3 per million euros traded when trading with the broker.†

### 3. Toxicity

In this paper, a trade is toxic over a given time window if the client can unwind the trade at a profit within the time window. Instead of classifying traders as informed or uninformed, the broker assesses the probability that each trade becomes toxic within a specified time window. Not all trades sent by better informed clients will be toxic, and not all trades sent by less informed clients will be benign. Thus, our models aim to predict price movements based on current features regardless of whether the trader is informed or not. Our methods, however, include the identity of the trader, so predicting toxicity of a trade will depend, among other features, on how often the client executed toxic trades in the past.

Denote time by  $t \in \mathcal{T} = [0, T]$ , where 0 is the start of the trading day and  $T$  is the end of the trading day. From this point forward, we use ‘exchange rate’ and ‘prices’ interchangeably. The best ask price and best bid price in the LOB of LMAX Exchange are denoted by  $(S_t^a)_{t \in \mathcal{T}}$  and  $(S_t^b)_{t \in \mathcal{T}}$ , respectively. Let  $\mathcal{G}$  be a toxicity horizon such that  $0 < \mathcal{G} \ll T$ , and let  $t \in$

$[0, T - \mathcal{G}]$ . We define the two stopping times

$$\begin{aligned} \tau_t^+ &= \inf \{u \in [t, T] : S_u^b > S_t^a\} \quad \text{and} \\ \tau_t^- &= \inf \{u \in [t, T] : S_t^b > S_u^a\}, \end{aligned}$$

with the convention that  $\inf \emptyset = \infty$ . The stopping time  $\tau_t^+$  is the first time after  $t$  that the best bid price is above the best ask price at time  $t$ . If  $\tau_t^+ < \infty$ , a buy trade executed at  $S_t^a$  becomes profitable for the client at time  $\tau_t^+$  before the end of the trading day because the client can unwind her position and collect the profit

$$S_{\tau_t^+}^b - S_t^a > 0. \quad (1)$$

Similarly,  $\tau_t^-$  is the first time after  $t$  that the best ask price is below the best bid price at time  $t$ . If  $\tau_t^- < \infty$ , a sell trade executed at  $S_t^b$  is profitable for the client at time  $\tau_t^-$  before the end of the trading day because the client can unwind her position and collect the profit

$$S_t^b - S_{\tau_t^-}^a > 0. \quad (2)$$

**DEFINITION 1 (Toxic trade)** Let  $\mathcal{G} > 0$  be a toxicity horizon. A client’s buy (resp. sell) filled by the broker at time  $t$  is toxic for the broker if  $\tau_t^+ \leq t + \mathcal{G}$  (resp. if  $\tau_t^- \leq t + \mathcal{G}$ ).

The above definition captures the broker’s exposure to adverse selection. A trade is labelled as toxic if over a given time window the client had the option to unwind the trade at a profit, in which case it would be a loss-leading trade for the broker. However, one cannot verify, ultimately, if the potentially toxic trade materialised as a loss to the broker or to another market participant. We do not have enough information to track each step, or potential step, in the life cycle of a trade to determine who made a loss or a gain — to make this assessment requires perfect knowledge of all trades by all market participants.‡

Figure 2 plots the trajectories of  $S_t^a$  and  $S_t^b$  for EUR/USD between 10:00:00 am and 10:00:10 am in LMAX Exchange on 28 June 2022. The dotted line is the best ask price and the dash-dotted line is the best bid price. The solid horizontal lines are the best ask price and the best bid price at 10:00:00 am.

In the figure, if a client buys from the broker at the best ask price at time  $t = 10:00:00$  am, then there is no opportunity for the trader to unwind the trade at a profit in the first ten seconds after the trade. However, had the trader sold to the broker at the best bid price at time  $t = 10:00:00$  am, then shortly after 10:00:01 the trade would be in-the-money for the client (i.e. toxic for the broker).

#### 3.1. Toxicity profiles

For a toxicity horizon  $\mathcal{G} > 0$ , we use both the client data and the LOB data to determine if the trades filled by LMAX Broker were toxic over the period  $\mathcal{G}$ . Table 2 shows the percentage of toxic trades executed by each client for  $\mathcal{G} \in \{1, 5, 10, 20, 30, 40, 50, 60, 70\}$  seconds.

‡ There are alternative ways of defining toxic flow with our proprietary data set. For instance, the broker values a trade after unwinding it with the first trade in the opposite direction, i.e. a first-in-first-out inventory valuation.

† In practice, not all clients receive the same bid and ask quotes.

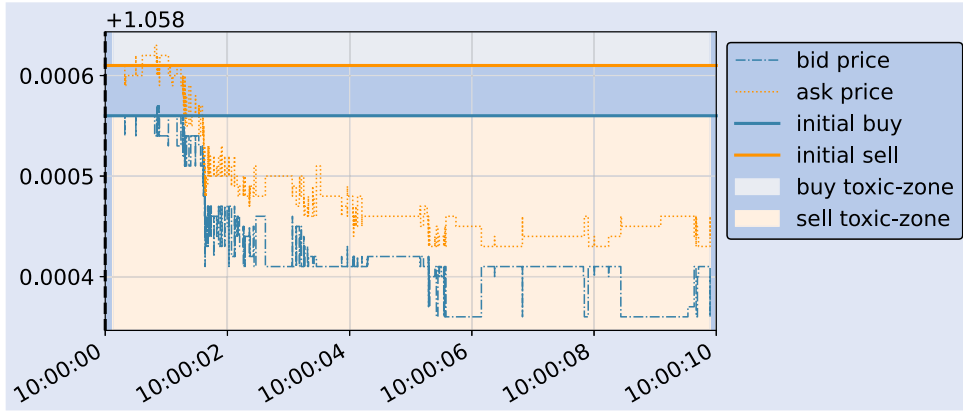


Figure 2. A client's sell trade that becomes toxic for the broker after a few seconds of filling the trade. The x-axis is time and the y-axis is in units of USD.

Table 2. Proportion of toxic trades (in %) between 28 June 2022 and 21 October 2022.

	toxicity horizon $\mathcal{G}$ in seconds								
	1	5	10	20	30	40	50	60	70
Client 1	6.7	25.7	38.4	51.5	58.7	63.4	66.7	69.3	71.3
Client 2	7.0	28.6	42.4	56.1	63.0	67.5	70.7	73.1	74.9
Client 3	7.0	26.0	38.6	51.1	58.2	62.6	65.8	68.2	70.4
Client 4	3.4	18.5	30.7	44.3	52.1	56.8	60.6	63.5	66.0
Client 5	8.3	22.1	31.7	42.9	50.1	55.4	59.6	62.2	63.9
Client 6	5.9	26.4	40.2	53.3	60.9	65.7	69.0	71.8	73.9
All clients	6.6	25.5	38.2	51.2	58.4	63.1	66.5	69.0	71.0

As expected, for short toxicity horizons only a small proportion of trades are toxic (e.g. 6.6% for a one second horizon), but as the toxicity horizon increases, the proportion of toxic trades grows considerably (e.g. it is roughly 70% after one minute). A simple mathematical argument can help us justify what we observe in the data. Consider a trader who sends a liquidity taking trade to the broker when the spread in the market is  $s$  and suppose that the profitability of unwinding the trade, which is  $-s$  at time zero, diffuses according to a scaled Brownian motion  $\sigma W_t$  with  $\sigma > 0$ . From the reflection principle, the probability that such a trade becomes toxic at any point between zero and  $\mathcal{G}$  seconds is given by

$$\begin{aligned}
 & \mathbb{P} \left( \sup_{t \in [0, \mathcal{G}]} \sigma W_t \geq s \right) \\
 &= \mathbb{P} \left( \sup_{t \in [0, \mathcal{G}]} W_t \geq \frac{s}{\sigma} \right) = 2\mathbb{P} \left( W_{\mathcal{G}} \geq \frac{s}{\sigma} \right) \\
 &= 2 \left( 1 - \Phi \left( \frac{s}{\sigma \sqrt{\mathcal{G}}} \right) \right), \tag{3}
 \end{aligned}$$

where  $\Phi$  is the standard normal cumulative distribution function. As the horizon  $\mathcal{G} \rightarrow \infty$ , the probability that the trade is toxic at some point converges to one. This does not mean that the broker loses money on each trade. That would happen only if the broker failed to hedge in the lit market and every liquidity unwound exited at the first profitable moment. In reality, the broker's inventory shifts constantly due to both incoming trades and her own hedging in the lit market.

Arguably, a client can be labelled as toxic according to the percentage of their trades that are in-the-money after a set time

frame, e.g. after  $\mathcal{G}$  in Definition 1 above. Figure 3 shows the toxicity profiles of two clients trading EUR/USD on 8 July 2022 with LMAX Broker. The figures summarise all trades by clients A and B as follows. For each trade on 8 July 2022, the plot shows the profitability, from the client's perspective, of unwinding each trade a given number of seconds after the trade. Here, the x-axis goes from zero to ten seconds and the y-axis is in dollars per million euros traded.

From figure 3, and all else being equal, a broker would prefer to provide liquidity to client B instead of client A. More than 50% of the times that client A trades, the broker is exposed to making a loss on the trade in less than half a second. On the other hand, the median trajectory of profitability (blue line) for client B is below zero.†

### 3.2. Features to predict toxic trades

For each client, the broker uses features that reflect (i) the state of the LOB, (ii) recent activity in the LOB, and (iii) the cash and inventory of the client in the EUR/USD currency pair.‡ Our features build on Ait-Sahalia *et al.* (2022), who propose using three different clocks to aggregate LOB data. Here, we compute eight LOB statistics, each with three clocks (time, volume, and transaction), and seven backward-looking *clock*

† Clearly, if the broker uses the flow from client A to inform other decisions and other investment strategies that are profitable for the broker, the net effect of trading at a loss with client A could be positive for the broker; see Cartea and Sánchez-Betancourt (2025).

‡ We use the transactions with the broker to compute the cash and inventory of the clients. Clients can trade elsewhere but this is unknown to the broker.

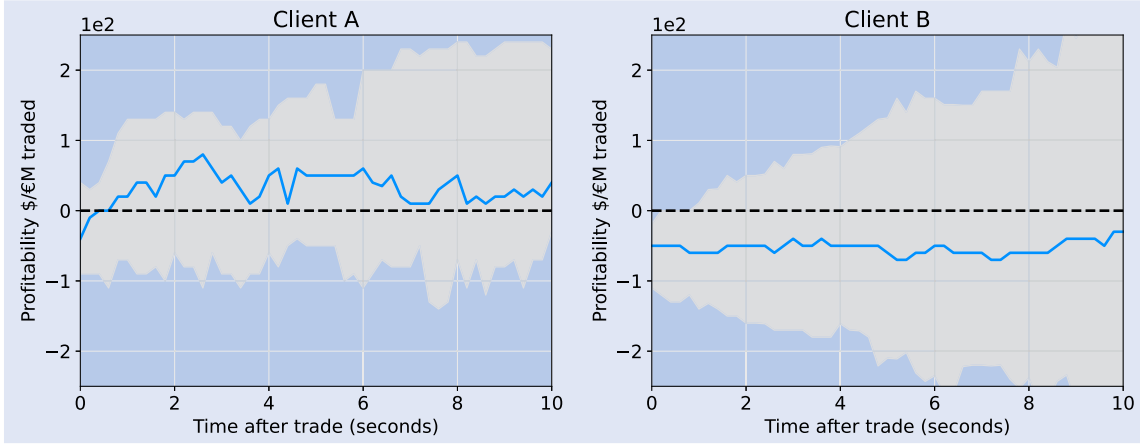


Figure 3. Profitability in dollars per million euros traded after a trade is executed. Panels correspond to two different clients. Blue line is the median trajectory and grey region is the 90% trajectory region. The  $x$ -axis is time and the  $y$ -axis is the profitability from the point of view of the client.

intervals of increasing length; thus, we have  $8 \times 3 \times 7 = 168$  clock-based features. We also employ fifteen other features (e.g. cash and inventory), so there are 183 features per client.

For each clock and a given interval in the past (as measured by the clock), we employ the following eight features: (a) volatility of the midprice, (b) number of trades executed by the client in the interval, (c) number of updates in the best quotes of LMAX Exchange in the interval, (d) return of the midprice over the interval, (e) average transformed volume in the best bid price, (f) average transformed volume in the best ask price, (g) average spread, and (h) average imbalance of the best available volumes, where “average” is simple arithmetic mean within an interval.

The three clocks provide alternative ways of deciding the datapoints that fall in a time interval. For example, with a time clock, the spread over the last two seconds is computed with all datapoints in the last 2 seconds. With a volume clock, the spread over the last  $V$  units of volume traded is computed with all the data in the past (chronologically) until we gather  $V$  units of volume traded. With a transaction clock, the spread over the last  $k$  transactions is computed with the datapoints from the last  $k$  transactions. See Appendix 1 for more details.

For a given client  $c \in \mathcal{A}$ , we use the following additional features: (i) cash, (ii) inventory, (iii) volume of the order, (iv) spread in the market just before the order arrives, (v) imbalance in the LOB just before the order arrives, (vi) transformed available volume in the best bid, (vii) transformed available volume in the best ask, (viii) last ask price at the time of trade, (ix) last bid price at the time of trade, (x) last midprice at the time of trade, (xi) total number of market updates since starting date, (xii) number of trades made by client  $c$ , (xiii) total number of trades executed by all clients, (xiv) volatility estimate of the mid-price, and (xv) proportion of previous toxic trades executed by client  $c$ . We apply log-transformations to stabilise scale and limit outlier leverage: a signed log-transformation of the form  $\text{sign}(x) \log(1 + |x|)$  for (i) cash and (ii) inventory, and a  $\log(1 + x)$  transform for volumes in (iii), (vi), and (vii). These monotone transforms reduce heteroskedasticity and outlier influence and improve numerical conditioning for the downstream models; see e.g. West (2022). The fifteen features above account for both the

state of the LOB, and the cash and inventory of the client. The remaining 168 features account for recent activity in the LOB. These are features (a)–(h) above, measured for each of the seven intervals and for each of the three clocks to obtain a total of  $8 \times 7 \times 3 = 168$  features. Thus, for each client, we employ  $15 + 8 \times 7 \times 3 = 183$  features to predict the toxicity of their trades.

In Appendix A.3 we illustrate that the performance of PULSE does not improve with additional information from the volume and transaction clock. In what follows we use the features provided by the three clocks to align with prior work and for completeness.

#### 4. The PULSE method

For a trade from client  $c \in \mathcal{A}$  filled at time  $t$ , let  $\mathbf{x}_t \in \mathbb{R}^M$  denote the features observed at  $t$  (with  $M$  the number of features). Define  $y_t \in \{0, 1\}$  as the toxicity indicator evaluated  $\mathcal{G} > 0$  after  $t$ , with  $y_t = 1$  if the trade is toxic within the horizon, and 0 otherwise. Thus,  $\mathbf{x}_t$  is observed at time  $t$ , whereas  $y_t$  is known at time  $t + \mathcal{G}$ .

For each trading day we collect observations  $(\mathcal{D}_i)_{i \in I}$ , where  $\mathcal{D}_i = (\mathbf{x}_i, y_i)$ ,  $I = \{1, 2, \dots, N\}$ ,  $t_i$  is the arrival time of observation  $i$ , and  $N$  is the number of trades that day. For notational convenience we write  $\mathcal{D}_i$  for  $\mathcal{D}_i$  and refer to the dataset as  $(\mathcal{D}_i)_{i \in I}$ . For  $n \in \mathbb{N}$ , let  $\mathcal{D}_{1:n} = \{\mathcal{D}_1, \dots, \mathcal{D}_n\}$  denote the first  $n$  observations; in particular, for  $n \leq N$  we have  $\mathcal{D}_{1:n} \subseteq \mathcal{D}_{1:N}$ .

Conditioned on the features  $\mathbf{x}_t$ , the toxicity indicator  $y_t \in \{0, 1\}$  is modelled as Bernoulli random variable with probability mass function

$$p(y | \boldsymbol{\theta}; \mathbf{x}_t) = \text{Bern}(y | \sigma(\mathbf{w}^\top g(\boldsymbol{\psi}; \mathbf{x}_t))), \quad y \in \{0, 1\}, \quad (4)$$

where  $g : \mathbb{R}^M \rightarrow \mathbb{R}^L$  is the output-layer of a NNNet,  $\boldsymbol{\theta} = (\mathbf{w}, \boldsymbol{\psi})$  are the parameters of the neural network,  $\mathbf{w}$  are the parameters of the *last layer*, and  $\boldsymbol{\psi}$  are the parameters in the *hidden layers*. Here, and throughout the paper, we adopt the convention that  $p$  denotes a likelihood function or a posterior

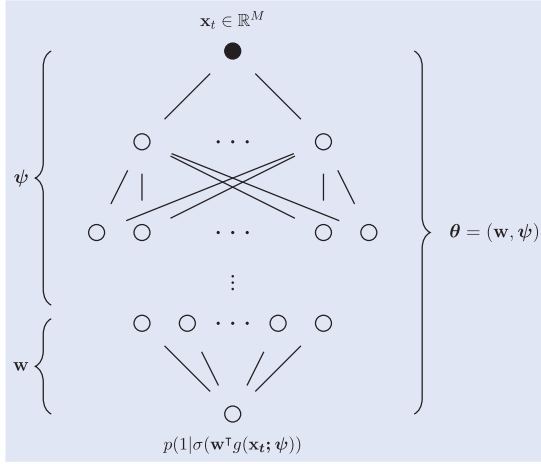


Figure 4. PULSE architecture for an MLP. The MLP is parameterised by  $\theta = (\boldsymbol{\psi}, \mathbf{w})$ , where  $\boldsymbol{\psi}$  are the parameters in the hidden layers and  $\mathbf{w}$  are the parameters in the last layer.

density function. The function  $\text{Bern}(\cdot | \cdot)$  is given by

$$\text{Bern}(a | b) = b^a (1 - b)^{1-a}, \quad a \in \{0, 1\} \text{ and } b \in (0, 1).$$

We refer to  $\mathbf{w} \in \mathbb{R}^L$  as the last-layer parameters and we refer to  $\boldsymbol{\psi} \in \mathbb{R}^D$  as the feature-transform parameters. The function  $\sigma(\mathbf{w}^\top g(\boldsymbol{\psi}; \mathbf{x}_t))$  is a NNet for classification where  $\sigma(x) = (1 + \exp(-x))^{-1}$  is the sigmoid function. Figure 4 shows a graphical representation of the parameters that PULSE updates when the NNet is a multilayered-perceptron (MLP). Although we choose an MLP in the experiments, PULSE can be used with any NNet architecture.

#### 4.1. Sequential learning

We update the model parameters sequentially after each observed toxicity label to incorporate new information. Specifically, we update  $\theta = (\mathbf{w}, \boldsymbol{\psi})$  in (4) after each new  $y_t$  is observed, i.e. after observing if the trade is toxic. In practice, the dimension  $D$  of the featured-transformed parameters and the dimension  $M$  of the feature space satisfy  $D \gg M$ , so it is costly to update  $\theta$  after each new observation using standard training techniques. Thus, the literature proposes various approaches to estimate the parameters of a NNet at a lower computational cost. In this paper, we build on two of these methods: lottery-ticket and last-layer methods.

Lottery-ticket methods exploit the over-parametrisation of NNets, in the sense that ‘a randomly-initialised dense NNet contains subnetworks that, when trained in isolation, reach test accuracy comparable to that of the original network’, (Frankle and Carbin 2019). The lottery-ticket hypothesis states that such a subnetwork exists, and subnetworks satisfying the lottery-ticket hypothesis are called *winning tickets*. *Winning tickets* are linear projections of the NNet parameters onto a subspace; see Li *et al.* (2018) and Larsen *et al.* (2022). Duran-Martin *et al.* (2022) use the lottery-ticket hypothesis with a relatively small linear subspace, and use the extended Kalman filter (EKF) algorithm to propose a sequential update of the subspace parameters.

Alternatively, last-layer methods pre-train the NNet parameters in a *warmup* phase and then perform sequential updates on the last-layer parameters, see Murphy (2023, S. 17.3.5). Here, we employ both methods. Specifically, we propose a Bayesian approach to update the parameters of a NNet sequentially for both the subspace parameters in the hidden layers and all of the parameters in the last layer. While previous literature focuses on updating either the last-layer or all parameters when performing online learning, ours is the first work that projects the parameters of the hidden layer and updates all of the units in the last layer. This decomposition enables full-rank updates of the last-layer parameters while restricting hidden-layer updates to a linear subspace. In doing so, it balances the rapid sequential updates typical of subspace neural networks with the statistical efficiency characteristic of last-layer methods.

In particular, we modify (4) and decompose the parameters of the hidden layer  $\boldsymbol{\psi} \in \mathbb{R}^D$  as an affine projection of the form

$$\boldsymbol{\psi} = \mathbf{A}\mathbf{z} + \mathbf{b}, \quad (5)$$

where  $\mathbf{A} \in \mathbb{R}^{D \times d}$  is the fixed projection matrix and  $\mathbf{z} \in \mathbb{R}^d$  are the projected (subspace) parameters such that  $d \ll D$ , and  $\mathbf{b} \in \mathbb{R}^D$  is the offset term. The decomposition (5) provides a linear-subspace formulation of the lottery-ticket hypothesis and enables efficient updates of the hidden layer parameters. With this projection, we rewrite (4) as  $p(y | \mathbf{z}, \mathbf{w}; \mathbf{x}_t) = \text{Bern}(y | \sigma(\mathbf{w}^\top g(\mathbf{A}\mathbf{z} + \mathbf{b}; \mathbf{x}_t)))$ ,  $y \in \{0, 1\}$ . To simplify notation, we define  $h(\mathbf{z}; \mathbf{x}_t) = g(\mathbf{A}\mathbf{z} + \mathbf{b}; \mathbf{x}_t)$  and write

$$p(y | \mathbf{z}, \mathbf{w}; \mathbf{x}_t) = \text{Bern}(y | \sigma(\mathbf{w}^\top h(\mathbf{z}; \mathbf{x}_t))). \quad (6)$$

Our procedure has two stages: (i) an offline warmup phase that estimates  $\mathbf{A}$  and  $\mathbf{b}$  and uses  $\mathcal{D}_{\text{warmup}}$  to select hyperparameters and (ii) an online phase that performs fast sequential updates of  $(\mathbf{w}, \mathbf{z})$  on the live stream  $\mathcal{D}_{\text{deploy}}$  while holding  $\mathbf{A}, \mathbf{b}$  fixed. Here,  $\mathcal{D}_{\text{warmup}}$  is a one-off historical window, whereas  $\mathcal{D}_{\text{deploy}}$  grows over time and is used for online learning and evaluation.

Figure 5 illustrates the workflow: an offline warmup phase using  $\mathcal{D}_{\text{warmup}}$  (Section 4.1.1) followed by an online deploy phase using  $\mathcal{D}_{\text{deploy}}$  (Section 4.1.2). The deploy phase proceeds indefinitely as trades arrive (the time index  $T$  may grow over time).

**4.1.1. Warmup phase: estimating the projection matrix and the offset term.** This phase estimates  $\mathbf{A}$  and  $\mathbf{b}$ , and assigns prior distributions for  $\mathbf{w}$  and  $\mathbf{z}$ . Given the size of the dataset, we divide  $\mathcal{D}_{\text{warmup}}$  into  $B$  non-intersecting random batches  $\mathcal{D}_{(1)}, \dots, \mathcal{D}_{(B)}$  such that

$$\bigcup_{b=1}^B \mathcal{D}_{(b)} = \mathcal{D}_{\text{warmup}}.$$

To estimate  $\mathbf{b}$  and  $\mathbf{A}$ , we use mini-batch stochastic gradient descent (SGD) over  $\mathcal{D}_{\text{warmup}}$  to minimise the negative loss-function

$$-\log p(\mathcal{D} | \boldsymbol{\theta}) = -\sum_{n=1}^N \log p(y_n | \boldsymbol{\theta}, \mathbf{x}_n), \quad (7)$$

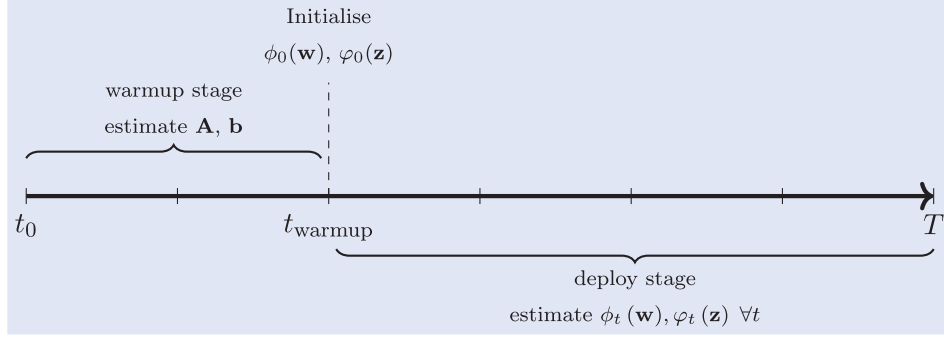


Figure 5. Warmup and deployment stages. We use all data available from  $t_0$  to  $t_{\text{warmup}}$  to estimate  $\mathbf{A}$  and  $\mathbf{b}$ . At  $t_{\text{warmup}}$ , we initialise the variational approximations  $\phi_0(\mathbf{w})$  and  $\varphi_0(\mathbf{z})$ . Finally, for  $t > t_{\text{warmup}}$ , we estimate  $\mathbf{w}_t$  and  $\mathbf{z}_t$ .

where  $\mathcal{D}$  is any random batch. The vector  $\mathbf{b}$  is given by

$$\mathbf{b} = \arg \min_{\boldsymbol{\theta}} -\log p(\mathcal{D} | \boldsymbol{\theta}). \quad (8)$$

Singular-value decomposition (SVD) over the iterates of the SGD optimisation procedure gives the projection matrix  $\mathbf{A}$ . To avoid redundancy, we skip the first  $n$  iterations and store the iterates every  $k$  steps. The dimension of the subspace  $d$  is found via hyperparameter tuning and convergence to a local minimum of (7) is obtained through multiple passes of the data. Algorithm 1 shows the training procedure for a number  $E$  of epochs.

---

**Algorithm 1:** MAP parameter estimation via batch SGD

---

```

1 def warmupParameters:
2 Initialise model parameters  $\boldsymbol{\theta} = (\boldsymbol{\psi}, \mathbf{w})$ 
3 foreach epoch  $e = 1, \dots, E$  do
4     foreach batch  $m = 1, \dots, M$  do
5         gradient =  $-\nabla_{\boldsymbol{\theta}} \log p(\mathcal{D}_{(m)} | \boldsymbol{\theta})$ 
6          $\boldsymbol{\theta}^{(e)} \leftarrow \boldsymbol{\theta}^{(e-1)} - \alpha \boldsymbol{\kappa}(\text{gradient})$ 
    
```

---

In Algorithm 1, the parameter  $\alpha$  is the learning rate, and the function  $\boldsymbol{\kappa} : \mathbb{R}^M \rightarrow \mathbb{R}^M$  is the per-step transformation of the Adam algorithm; see Kingma and Ba (2015). At the end of the  $E$  epochs, we obtain  $\boldsymbol{\theta}^{(E)} = (\boldsymbol{\psi}^{(E)}, \mathbf{w}^{(E)})$ . Then, the offset term  $\mathbf{b}$  is given by

$$\mathbf{b} = \boldsymbol{\psi}^{(E)},$$

and we stack the history of the SGD iterates. To avoid redundancy, we skip the first  $n$  iterates of the SGD, which are stored at every  $k$  steps. We let

$$\mathcal{E} = \begin{bmatrix} \text{---} & \boldsymbol{\psi}^{(n)} & \text{---} \\ \text{---} & \boldsymbol{\psi}^{(n+k)} & \text{---} \\ \text{---} & \boldsymbol{\psi}^{(n+2k)} & \text{---} \\ & \vdots & \\ \text{---} & \boldsymbol{\psi}^{(E)} & \text{---} \end{bmatrix} \in \mathbb{R}^{\hat{E} \times D},$$

where  $\hat{E} = E - n + 1$ . With the SVD decomposition  $\mathcal{E} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}$  and the first  $d$  columns of the matrix  $\mathbf{V}$ , the projection

matrix is

$$\mathbf{A} = \begin{bmatrix} | & | & \dots & | \\ \mathbf{V}_1 & \mathbf{V}_2 & \dots & \mathbf{V}_d \\ | & | & \dots & | \end{bmatrix},$$

where  $\mathbf{V}_k$  denotes the  $k$ th column of  $\mathbf{V}$ .

#### 4.1.2. Deploy phase: online estimation of last-layer and subspace-feature-transform parameters.

Here, we derive a novel, sample-free, and closed-form update rule that estimates the parameters of (6) sequentially. Specifically, in Proposition A.1 in the Appendix, we find a set of fixed-point equations for the update rule. These equations need many iterations to converge and are computationally inefficient. Thus, Corollary A.1 computes the gradient with respect to the subspace parameters to simplify the computations. Finally, Theorem 2 uses a Taylor expansion of the measurement model to obtain a closed-form solution to the set of fixed-point equations. This is computationally efficient because the update can be obtained in a single iteration.

We introduce Gaussian priors for both  $\mathbf{w}$  and  $\mathbf{z}$  at the beginning of the deploy stage. Let  $n = 0$  denote the last timestamp in the warmup dataset and  $n = 1$  the first timestamp of the deploy dataset. Denote by  $\phi_n$  and  $\varphi_n$  the posterior distribution estimates for  $\mathbf{w}$  and  $\mathbf{z}$  at time  $t$ , respectively. The initial estimates are given by

$$\begin{aligned} \phi_0(\mathbf{w}) &= \mathcal{N}(\mathbf{w} | \mathbf{w}^{(M)}, \sigma_{\mathbf{w}}^2 \mathbf{I}), \\ \varphi_0(\mathbf{z}) &= \mathcal{N}(\mathbf{z} | \boldsymbol{\psi}^{(M)} \mathbf{A}, \sigma_{\mathbf{z}}^2 \mathbf{I}), \end{aligned}$$

where  $(\mathbf{w}^{(M)}, \boldsymbol{\psi}^{(M)})$  are the last iterates in the warmup stage,  $\sigma_{\mathbf{w}}^2$  and  $\sigma_{\mathbf{z}}^2$  are the coefficients of the prior covariance matrix,  $\mathbf{I}$  is the identity matrix, and recall that  $\mathbf{A}$  is the projection matrix.

For  $n \geq 1$ , the variational posterior estimates are given by

$$\phi_n(\mathbf{w}) = \mathcal{N}(\mathbf{w} | \mathbf{v}_n, \boldsymbol{\Sigma}_n) \quad \text{and} \quad \varphi_n(\mathbf{z}) = \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}_n, \boldsymbol{\Gamma}_n).$$

Next, to find the posterior parameters  $\boldsymbol{\mu}_n, \mathbf{v}_n, \boldsymbol{\Gamma}_n, \boldsymbol{\Sigma}_n$ , we recursively solve the following variational inference (VI) optimisation problem

$$\begin{aligned} &\boldsymbol{\mu}_n, \mathbf{v}_n, \boldsymbol{\Gamma}_n, \boldsymbol{\Sigma}_n \\ &= \arg \min_{\boldsymbol{\mu}, \mathbf{v}, \boldsymbol{\Gamma}, \boldsymbol{\Sigma}} \text{KL}(\mathcal{N}(\mathbf{w} | \mathbf{v}, \boldsymbol{\Sigma}) \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \boldsymbol{\Gamma})) \end{aligned}$$

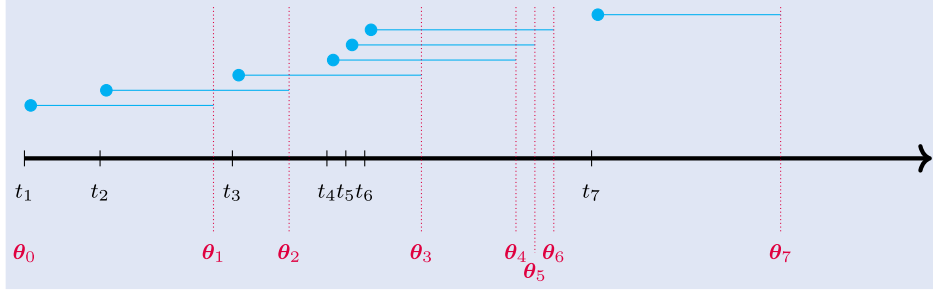


Figure 6. Asynchronous predict-update steps: trades arrive at irregular times  $\{t_i\}_i$ . An update to the model is only possible if  $t_{i+1} > t_i + \mathcal{G}$ , when we know whether the trade was toxic or benign. In this example, the model parameters  $\theta_0$  are known at time  $t_1$ , when a new trade arrives. When a second trade arrives, at time  $t_2$ , we do not know whether the previous trade was toxic or benign at  $t_1$ , so we use the model weights  $\theta_0$  to make a prediction. The next trade arrives at time  $t_3 > t_1 + \mathcal{G}$ , so we use  $\theta_1$  to make a prediction. Finally, multiple trades arrive consecutively at times  $t_4$ ,  $t_5$ , and  $t_6$ , in which case we use  $\theta_2$ . The last trade to arrive at time  $t_7$  uses  $\theta_6$ . In this example  $\theta_3$ ,  $\theta_4$ , and  $\theta_5$  were never used to make a prediction because trades did not arrive during the period  $[t_3 + \mathcal{G}, t_6 + \mathcal{G}]$ .

$$\|\phi_{n-1}(\mathbf{w})\phi_{n-1}(\mathbf{z})p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)\|, \quad (9)$$

where KL is the Kullback–Leibler divergence

$$\text{KL}(p(x) \| q(x)) = \int p(x) \log \left( \frac{p(x)}{q(x)} \right) dx,$$

for probability density functions  $p$  and  $q$  with the same support. The optimisation in (9) generalises the update rule for the Kalman filter when the parameters do not have a drift; see e.g. Lambert *et al.* (2021). The following theorem shows the update and prediction equations of the PULSE method.

**THEOREM 2 (PULSE)** *Suppose  $\log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)$  is differentiable with respect to  $(\mathbf{z}, \mathbf{w})$  and the observations  $\{y_n\}_{n=1}^N$  are conditionally independent over  $(\mathbf{z}, \mathbf{w})$ . Write the mean of the target variable  $y_n$  as a first-order approximation of the parameters centred around their previous estimate. Let  $\sigma(x) = (1 + \exp(-x))^{-1}$  be the sigmoid function and  $\sigma'(x) = \sigma(x)(1 - \sigma(x))$  its derivative. Then, an approximate solution to (9) is given by*

$$\mathbf{v}_n = \mathbf{v}_{n-1} + \Sigma_{n-1} h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n) (y_n - \sigma(\mathbf{v}_{n-1}^\top h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n))), \quad (10)$$

$$\Sigma_n^{-1} = \Sigma_{n-1}^{-1} + \sigma'(\mathbf{v}_{n-1}^\top h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n)) h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n)^\top h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n), \quad (11)$$

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{n-1} + \Gamma_{n-1} \nabla_{\mathbf{z}} h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n) \times (y_n - \sigma(\mathbf{v}_{n-1}^\top h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n))), \quad (12)$$

$$\Gamma_n^{-1} = \Gamma_{n-1}^{-1} + \sigma'(\mathbf{v}_{n-1}^\top h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n)) \times \nabla_{\mathbf{z}} h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n) \nabla_{\mathbf{z}} h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n)^\top, \quad (13)$$

where  $\boldsymbol{\mu}_n$ ,  $\Gamma_n$  are the estimated mean and covariance of the projected-hidden-layer parameters at step  $n$ , and  $\mathbf{v}_n$ ,  $\Sigma_n$  are the estimated mean and covariance matrix of the last-layer parameters.

As a corollary, a variant of PULSE can be derived to model any other member of the exponential family by replacing the mean and covariance of the target distribution of choice. See Appendix 2 for a proof of Theorem 2.

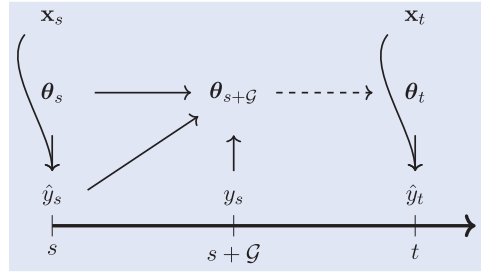


Figure 7. PULSE update procedure. For simplicity, we take  $s + \mathcal{G} < t$ .

## 5. Asynchronous learning and decision making

Next, we discuss how we deploy and evaluate the performance of the online PULSE methodology with asynchronous data. In classical filtering problems, as soon as new information arrives at, say, time  $t_i$ , the parameters  $\theta_{t_i}$  of the model are updated. Next, when a new trade arrives at time  $t_{i+1}$ , one uses the parameters  $\theta_{t_i}$  to estimate if a trade will be toxic. In our setting, however, an update at time  $t_{i+1}$  with  $\theta_{t_i}$  is only possible if  $t_{i+1}$  is greater than the time of last trade  $t_i$  plus the toxicity horizon  $\mathcal{G} > 0$ , i.e.  $t_{i+1} > t_i + \mathcal{G}$ . Otherwise, we use  $\theta_{t_j}$  to predict the probability of a toxic trade, with  $j = \arg \max_k t_{i+1} > t_k + \mathcal{G}$ . Figure 6 illustrates this procedure.

We employ the asynchronous online updating for PULSE and MLE and select hyperparameters over the warmup stage. Figure 7 shows how PULSE updates model parameters based on: current model parameters  $\theta$ , features  $\mathbf{x}$ , and outcome  $y$ . Here,  $\theta = (\boldsymbol{\psi}, \mathbf{w})$  are the model parameters one uses to produce  $p(y = 1 | \mathbf{x}, \theta)$ , with which we compute the prediction  $\hat{y}$ . We employ the predictions  $\hat{y}$  and the outcomes  $y$  to compute the accuracy defined above.

### 5.1. Model for decision making

Here, we devise brokerage strategies that employ predictions of toxic flow. We introduce a one-shot optimisation problem for the broker's strategy to internalise-externalise trades.

For method  $M \in \{\text{PULSE}, \text{LogR}, \text{RF}, \text{MLE}\}$ , let  $p^{+,M} \in (0, 1)$ , denote the probability that a buy order will be toxic and let  $p^{-,M}$  denote the probability that a sell order will be

toxic. Note that  $p^{+,M} + p^{-,M}$  does not necessarily add to 1. Let  $S/2 > 0$  denote the half bid-ask spread and let  $\eta > 0$  denote the shock to the midprice  $S$  if the trade is toxic; here we assume that  $\eta \in (S, \infty)$ . The broker controls  $\delta^\pm \in \{0, 1\}$ . When  $\delta^\pm = 0$  the broker externalises the trade and when  $\delta^\pm = 1$  the broker internalises the trade. The inventory of the broker is  $Q \in \mathbb{R}$ ; when  $Q > 0$  the broker is long and when  $Q < 0$  the broker is short. Assume all trades are for one unit of the asset. Then, the broker solves

$$\delta^{\pm*} = \arg \max_{\delta^\pm \in \{0, 1\}} \mathbb{E} \left[ \underbrace{\pm \delta^\pm (S \pm S/2)}_{\text{cash flow}} + \underbrace{(S \pm \eta Z) (Q \mp \delta^\pm)}_{\text{inventory valuation}} \right] - \underbrace{\phi (Q \mp \delta^\pm)^2}_{\text{inventory penalty}}, \quad (14)$$

where  $Z$  is a Bernoulli random variable with parameter  $p^{\pm, M}$ , and  $\phi \geq 0$  is an inventory penalty parameter. Intuitively, the broker optimises the expected mark-to-market value of her portfolio after internalising the trade adjusted by a quadratic penalty on inventory. The solutions to (14) are

$$\begin{aligned} \delta^{\pm*} &= \mathbb{1} \left( \frac{S/2}{\eta} - \frac{\phi}{\eta} \pm \frac{2\phi}{\eta} Q > p^{\pm, M} \right) \\ &= \mathbb{1} (p \pm \Phi Q > p^{\pm, M}), \end{aligned} \quad (15)$$

where  $p := S/2\eta - \phi/\eta$  and  $\Phi := 2\phi/\eta$ . We call  $\Phi$  the inventory aversion parameter and we call  $p$  the cutoff probability. The strategy internalises trades when the prediction  $p^{\pm, M}$  is lower than the cutoff probability  $p$  adjusted by the inventory of the broker  $Q$  and the inventory aversion parameter  $\Phi$ . When either

$$p + \Phi Q = p^{+,M} \quad \text{or} \quad p - \Phi Q = p^{-,M},$$

the broker is indifferent between internalising or externalising the trade in the market; this happens with probability zero.

The model variables  $S$ ,  $\eta$ , and  $\phi$  in (14) are not calibrated in our experiments. Their role here is to motivate the broker's decision rule (15) in terms of the cutoff probability  $p$  and the inventory aversion parameter  $\Phi$ .

Next, we study the case  $\Phi = 0$  in more detail. In Appendix A.2 we explore the case  $\Phi > 0$ .

## 5.2. Internalise-externalise strategy

Motivated by the mathematical framework in Section 5.1, below we introduce a family of predictions of toxicity based on the cutoff probability  $p \in [0, 1]$ . In what follows, we ignore the permanent price impact of externalising trades, as this would require the formulation of a stochastic control problem.† Below, when deploying our strategies, the historical data do not change.

† For a series of recent related papers on broker-client stochastic control problems see Cartea and Sánchez-Betancourt (2025), Bergault and Sánchez-Betancourt (2025), Cartea *et al.* (2025), Donnelly and

**DEFINITION 3** (*p*-predicted toxic trade) Let  $p \in [0, 1]$  and  $p(y = 1 | \mathbf{x}_n, \theta)$  be the output of a classifier. A trade is predicted to be toxic with cutoff probability  $p$  if

$$p(y = 1 | \mathbf{x}_n, \theta) > p. \quad (16)$$

We store the decision of a toxic trade in the variable

$$\hat{y}_{t_n}^p = \mathbb{1}(p(y = 1 | \mathbf{x}_n, \theta) > p). \quad (17)$$

We are interested in the predictive performance of the models as we vary the value of  $p$ . To this end, let  $y_{t_n} \in \{0, 1\}$  denote if a trade executed at time  $t_n$  was toxic at time  $t_n + \mathcal{G}$  ( $y_{t_n} = 1$  if toxic and  $y_{t_n} = 0$  otherwise). We employ the true positive rate and the false positive rate, which we define below.

**DEFINITION 4** The true positive rate (TPR) of a sequence of trades  $\{y_{t_n}\}_{n=1}^N$  with predictions  $\{\hat{y}_{t_n}^p\}_{n=1}^N$  at a cutoff probability  $p$  is

$$\text{TPR}_p = \frac{\sum_{n=1}^N \mathbb{1}(y_{t_n} = \hat{y}_{t_n}^p) \cdot \mathbb{1}(y_{t_n} = 1)}{\sum_{n=1}^N \mathbb{1}(y_{t_n} = 1)}. \quad (18)$$

**DEFINITION 5** The false positive rate (FPR) of a sequence of trades  $\{y_{t_n}\}_{n=1}^N$  with predictions  $\{\hat{y}_{t_n}^p\}_{n=1}^N$  at a cutoff probability  $p$  is

$$\text{FPR}_p = \frac{\sum_{n=1}^N \mathbb{1}(y_{t_n} \neq \hat{y}_{t_n}^p) \cdot \mathbb{1}(y_{t_n} = 0)}{\sum_{n=1}^N \mathbb{1}(y_{t_n} = 0)}. \quad (19)$$

Each choice of  $p$  induces a pair of values  $(\text{FPR}_p, \text{TPR}_p)$ . The graph of  $p \rightarrow (\text{FPR}_p, \text{TPR}_p)$  is known as the Receiver Operating Characteristic (ROC). Figure 8 shows the daily ROC of the models in the deploy stage with toxicity horizon of 30s.

The area under an ROC curve, called AUC, is used in the machine learning literature to compare classifiers; see, e.g. Fawcett (2006). Intuitively, the AUC is a measure to quantify a classifier's ability to distinguish between toxic and benign trades.

## 6. Experiments

We employ the methodology developed in the previous section with the following configuration. The NNet for PULSE is an MLP with three hidden layers, 100 units in each layer, and ReLU activation function. The number of epochs  $E$

Li (2025) and Wu and Jaimungal (2024). In these papers, the externalisation activity of the broker has a permanent price impact in prices. There, trades are toxic because informed traders observe the stochastic drift of the asset price and exploit that information optimally. The problem then is to characterise the equilibrium between the broker and the informed traders (the uninformed trader is usually assumed to be non-strategic). One can also allow the broker to have flexibility in the skewing mechanism, as opposed to using bid and ask quotes that are aligned to those in the lit market, see e.g. Barzykin *et al.* (2022, 2023). Furthermore, the broker can employ filtering techniques to identify toxic trades. For example, if skewed quotes which are 'expensive' are hid by a client, this may convey information about the toxicity of the trade; see e.g. Aqsha *et al.* (2024) where the authors filter the trading speed of informed traders in search of the signal.

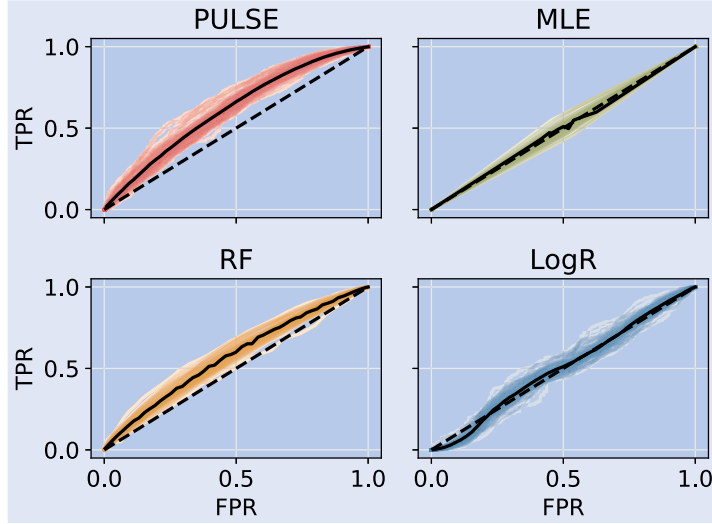


Figure 8. Daily ROC curves with toxicity horizon of 30s. We plot the daily ROC curve for each model at the end of each trading day. Each coloured line represents the ROC curve for a trading day. The solid black line is the average of the daily ROC curves. Finally, the black dashed line represents the ROC curve for a random classifier.

is 850, we skip the first 50 iterations of the optimisation procedure, the subspace dimension is  $d = 20$ , the learning rate is  $\alpha = 10^{-7}$ , and we store gradients every  $k = 4$  steps.† With this configuration, we estimate the 38, 700 units of the MLP during the warmup stage. For the deployment stage, PULSE updates 120 degrees of freedom; this accounts for less than half a percent of all parameters updated during the warmup stage. From a practical perspective, the memory cost of a single step of the algorithm (during the deployment stage) is  $O((L + D)^2)$ . In this paper, an update requires less than 1mb of memory because each unit is a 32bit float. Conversely, if we did not employ the subspace approach a single step would require 190gb of memory, making it infeasible to run on typical GPU devices.

We divide the dataset into a warmup dataset  $\mathcal{D}_{\text{warmup}}$  and a deploy dataset  $\mathcal{D}_{\text{deploy}}$ . Here,  $\mathcal{D}_{\text{warmup}}$  is from 28 June 2022 to 29 July 2022, and  $\mathcal{D}_{\text{deploy}}$  is from 1 August 2022 to 21 October 2022.

### 6.1. Benchmark models

We compare the performance of four methods: PULSE, logistic regression (LogR), random forests (RF), and a recursively updated maximum-likelihood estimator of a Bernoulli-distributed random variable (MLE). The MLE benchmark reflects a common industry practice in which toxicity is treated as a client-level attribute rather than a trade-specific one. We include LogR as a widely used linear baseline that helps quantify the added value of employing non-linear models such as neural networks. Finally, we include RF, a strong tree-based nonparametric method that is competitive on tabular data (e.g. McElfresh *et al.* 2023).

† Our results are robust to higher subspace dimensions. In particular, we find that using a dimension up to  $d = 1000$  does not yield a statistically significant outperformance over the baseline with  $d = 20$ .

With logistic regression, the probability that a trade is toxic is

$$p(y | \mathbf{w}_0; \mathbf{x}_t) = \text{Bern}(y | \sigma(\mathbf{w}_0^\top \mathbf{x}_t)), \quad y \in \{0, 1\}, \quad (20)$$

where  $\mathbf{w}_0$  is estimated maximising the log-likelihood with L-BFGS; see Liu and Nocedal (1989).

Next, RF is a bootstrap-aggregated collection of decorrelated trees. To predict if a trade is toxic one uses the average over the individual trees in the ensemble, see Section 15.1 of Hastie *et al.* (2001).

Further, we model the unconditional probability of a toxic trade as a Bernoulli-distributed random variable with mass function

$$p(y | \pi) = \text{Bern}(y | \pi), \quad y \in \{0, 1\}. \quad (21)$$

The maximum-likelihood estimator of the parameter  $\pi$ , given a collection  $\{z_1, \dots, z_N\}$  of Bernoulli-distributed samples, where  $z_n \in \{0, 1\}$ , is given by

$$\pi_{\text{MLE}} = \frac{1}{N} \sum_{n=1}^N \mathbb{1}(z_n = 1);$$

here,  $\mathbb{1}(\cdot)$  is the indicator function and we refer to this estimator as the MLE method. This quantity is updated after each new observation  $y_t$ .

The decision rule in (15) is directional and therefore induces two conditional problems: one for buy trades and one for sell trades. We consequently learn side-specific feature transformations and parameters. There are three implementations:

$$\mathbf{x}_{\text{bid}} \rightarrow \mathcal{M}_{\text{bid}} \rightarrow y, \quad \mathbf{x}_{\text{ask}} \rightarrow \mathcal{M}_{\text{ask}} \rightarrow y, \quad (22)$$

$$[\mathbf{x}_{\text{bid}}, \mathbf{x}_{\text{ask}}] \rightarrow \mathcal{M} \rightarrow [y_{\text{bid}}, y_{\text{ask}}], \quad (23)$$

$$[\mathbf{x}, \text{bid}/\text{ask}] \rightarrow \mathcal{M} \rightarrow y, \quad (24)$$

where  $\mathcal{M}$  is for model. We adopt (22) because it enables a direct comparison with the baselines (LogR, RF, MLE).

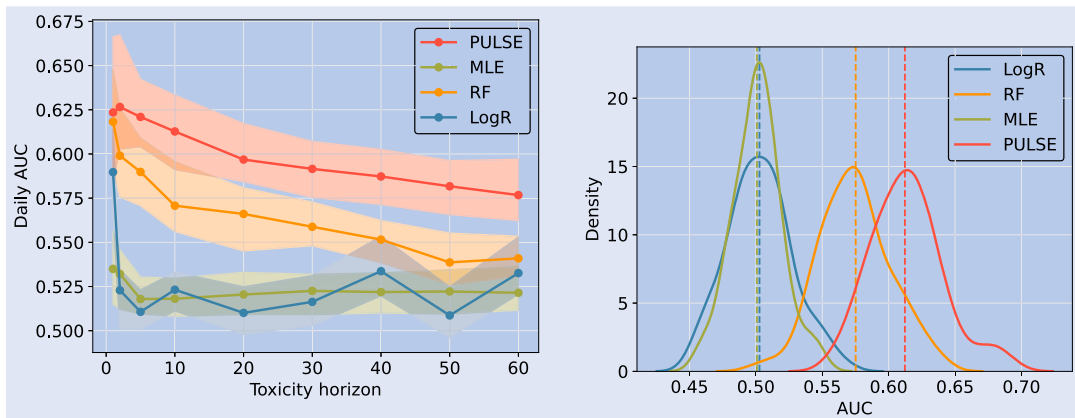


Figure 9. Left panel shows the median (line) and interquartile range (shaded) of daily AUC by toxicity horizon and model. Right panel shows the daily AUC of all models for a toxicity horizon of 30s. The dashed vertical lines correspond to the mean daily value. EUR/USD currency pair over the period 1 August 2022 to 21 October 2022.

In Appendix A.4 we study the performance when we train one model per client and use the client’s unique features.

*Online evaluation.* We evaluate MLE and PULSE online because both admit single-pass updates compatible with asynchronous labels. Our pipeline also supports online logistic regression; however, in our data, LogR underperforms even with batch retraining, yielding score distributions concentrated near the base rate, so online updates would not improve its behaviour. Fully online RF would require incremental trees or replay buffers which require more compute power takes longer to run. For robustness, in Appendix A.6 we re-train LogR and RF weekly on accrued data.

## 6.2. Model comparison

In this section, we analyse the AUC of the methods for various toxicity horizons. For each method, we compute the AUC for the sequence of trades of each day and show a density plot of such values in figure 9; recall that the deploy window is between 1 August 2022 and 21 October 2022.

PULSE has the highest average AUC among the four methods for all toxicity horizons. Furthermore, as the toxic horizon increases, the AUC of PULSE and RF decrease but the out-performance of PULSE over RF increases. MLE and LogR have the poorest performance. We observe that the AUC for PULSE and RF decreases as the toxicity horizon increases, reflecting that the prediction problem is more difficult over longer horizons. Empirically, the variance of profitability increases with time (see figure 3), which reduces the predictive power of these methods. For MLE, the AUC remains close to 0.5 across horizons, highlighting that it acts as a random classifier. On the other hand, LogR shows high variability across horizons, mainly because its linear structure is not well-suited to capture the dynamics of the data, leading to poor calibration and unstable performance.

Figure 10 shows the five-day exponentially-weighted average of the AUC for each day. RF, PULSE, and MLE attain their maximum values at the beginning of the deploy period and then decay over time, while PULSE maintains a steady performance because PULSE updates its parameters with each new observation. The performance of MLE and LogR

is similar. In our data, LogR does not find a useful linear boundary. Consequently, LogR’s predicted probabilities cluster near the unconditional rate and are highly correlated with MLE’s, yielding little time variation.

Figure 11 shows a five-day exponentially-weighted moving average (with decay 1/3) of the AUC for the various toxicity horizons across time.

We observe that as the toxicity horizon increases, the maximum AUC decreases almost uniformly.

## 6.3. Trade prediction effectiveness and missed opportunities

We employ data for the EUR/USD currency pair over the period 1 August 2022 to 21 October 2022 to test the internalisation-externalisation strategy. As above, we ignore inventory aversion (i.e.  $\Phi = 0$ ).

In what follows, we define the avoided profit of an externalised trade to be the PnL of that trade had it been internalised with unwinding at the end of the toxicity horizon. Figure 12 reports the PnL (y-axis) of the internalised trades and the avoided profit (x-axis) of the externalised trades when the broker uses the internalisation-externalisation strategy (15) and for PULSE, MLE, LogR, and RF. The points shown for each method and toxicity horizon are the ones that maximised PnL over all possible cutoff probabilities  $p \in \{0.05, 0.15, 0.25, \dots, 0.95\}$ .

The broker starts with zero inventory at the beginning of 1 August 2022 and she crosses the spread to unwind the internalised trades at the end of the toxicity horizon. We keep track of the PnL they would have obtained over the toxicity horizon.† The inventory is in euros (€) and the PnL is in dollars (\$). Each trade is for the median quantity which is €2,000 in our dataset. When a trade is toxic, the median loss to unwind the position is  $\$7 \times 10^{-5}$  per euro traded and when the trade is not toxic, the median profit is  $\$8 \times 10^{-5}$  per euro traded.

For each toxicity horizon, the internalisation-externalisation strategy informed by PULSE attains the

† Studying the variance of these results is more challenging and left for future research.

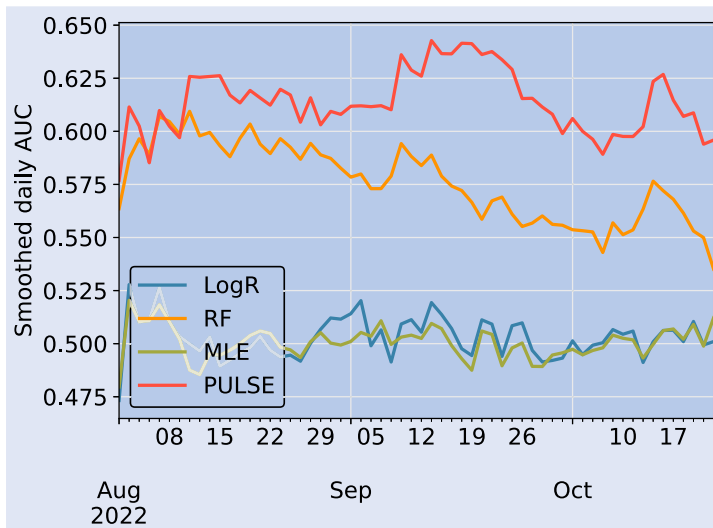


Figure 10. Five-day exponentially-weighted moving average of AUC over time. The toxicity horizon is 30 seconds.

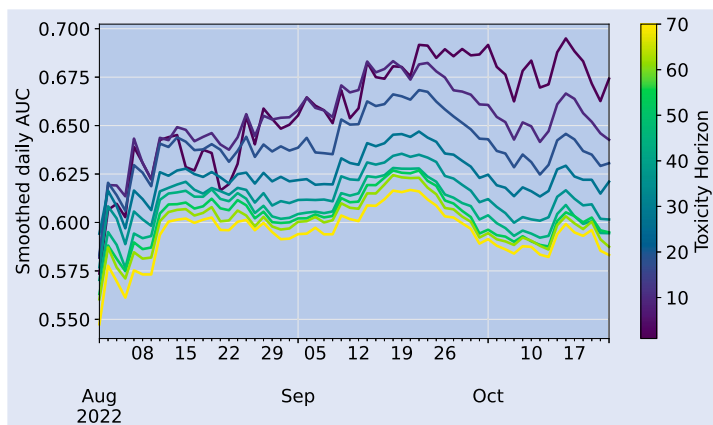


Figure 11. Five-day exponentially-weighted moving average of AUC over time for PULSE across toxicity horizons.

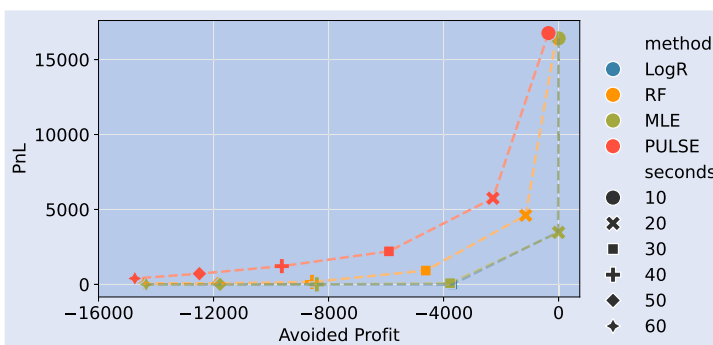


Figure 12. PnL and avoided profit for various toxicity horizons and for PULSE, MLE, LogR, and RF. Each point shows the highest possible PnL for a given method and toxicity horizon where the maximum is taken across cutoff probability where  $p \in \{0.05, 0.15, 0.25, \dots, 0.95\}$ .

highest PnL and the lowest avoided profits,<sup>†</sup> see red dots joined by the dash line. These results show the added economic value that one obtains when informing trading strategy (15) with the predictions made by PULSE. Indeed, we show that the higher quality of predictions obtained by PULSE produce higher PnL and lower avoided profits.

Next, figure 13 shows a histogram of  $p^{\pm, M}$  for  $M \in \{PULSE, MLE, LogR, RF\}$  and toxicity horizon of 10s and

<sup>†</sup> A negative avoided profit is an incurred loss had the trade been internalised and unwound at the toxicity horizon.

figure 14 shows how the percentage of internalised volume depends on the cutoff probability.

MLE produces an almost constant score (around the unconditional toxicity rate). For LogR, the concentration of probabilities below 0.5 is due to the nonlinear structure of the data. Consequently, when the cutoff crosses these concentrated values, the internalised volume changes abruptly. In contrast, PULSE produces a broader, more dispersed (higher-entropy) distribution of probabilities, enabling smoother and more informative policy adjustments.

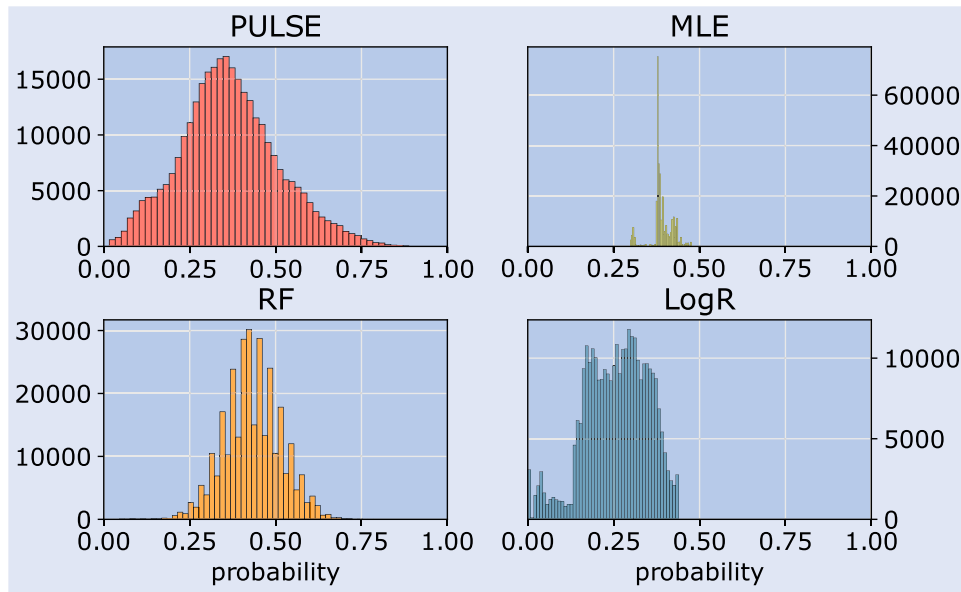


Figure 13. Histograms of predicted toxicity probabilities by model (pooled across clients). Toxicity horizon is ten seconds. EUR/USD currency pair over the period 1 August 2022 to 21 October 2022.

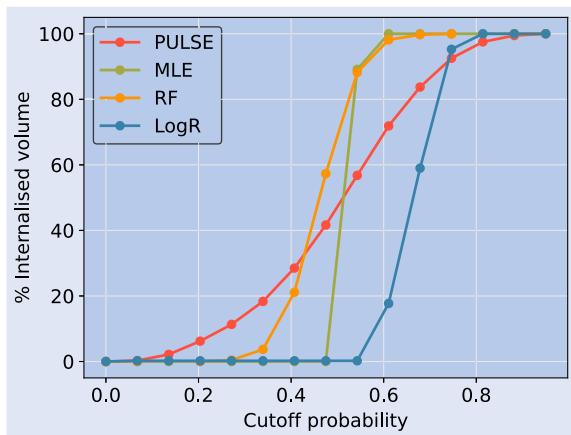


Figure 14. Percentage of internalised volume as a function of the cutoff probability  $p$  for toxicity horizon of 20s. EUR/USD currency pair over the period 1 August 2022 to 21 October 2022.

Figure 14 plots the percentage of internalised volume as a function of the cutoff  $p$ . Step-like moves occur when  $p$  crosses regions where scores are concentrated. For MLE, scores are almost constant at the unconditional toxicity rate (as seen in the previous histogram), thus, the percentage of internalised volume jumps when  $p$  crosses that level. For LogR, scores cluster near the base rate, producing a sharp transition similar to that of MLE. RF shows smaller but noticeable jumps, reflecting its narrower score dispersion. In contrast, PULSE produces a more dispersed score distribution and therefore a smoother curve. See Appendix A.8 for an analysis on precision and recall.

## 7. Conclusions

We employed machine learning and statistical methods to detect toxic flow. We also developed a novel method for online training of neural networks, which we call PULSE. We use PULSE to estimate the parameters of a neural network (sequentially) that computes the probability that an incoming

trade will be toxic. The out-of-sample performance of the multilayered-perceptron (MLP) trained with PULSE is high, stable, and outperforms the other methods we considered.

We proposed a broker’s strategy that uses these predictions to decide which trades are internalised and which are externalised by the broker. The mean PnL of the internalise-externalise strategy we obtain when training the MLP with PULSE is the highest (when compared with the benchmarks) and it is robust to model parameter choices. Future research will consider a hierarchical version of the problem, where there is a structure for toxicity common to all traders. The methodology proposed by PULSE can also be used in other areas of finance where sequential updates are desirable, such as in the prediction of fill-rate probabilities, and in multi-armed bandit problems for trading; see, e.g. Arroyo *et al.* (2024) and Cartea *et al.* (2023).

## Acknowledgements

We thank Andrew Stewart, Alistair Sturgiss, Fayçal Drissi, Patrick Chang, Álvaro Arroyo, Sergio Calvo, the participants at the Victoria Seminar Series, the Crossroads Seminar, and the University of Edinburgh for comments. ChatGPT suggested the name PULSE for our algorithm.

## Comments

For the purpose of open access, the authors have applied a CC BY public copyright licence to any author accepted manuscript version arising from this submission.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Álvaro Cartea  <http://orcid.org/0000-0002-7426-4645>

Gerardo Duran-Martin  <http://orcid.org/0009-0000-0764-4958>

Leandro Sánchez-Betancourt  <http://orcid.org/0000-0001-6447-7105>

## References

- Ait-Sahalia, Y., Fan, J., Xue, L. and Zhou, Y., How and when are high-frequency stock returns predictable? Technical Report, National Bureau of Economic Research, 2022.
- Amihud, Y. and Mendelson, H., Dealership market: Market-making with inventory. *J. Financ. Econ.*, 1980, **8**(1), 31–53.
- Aqsha, A., Drissi, F. and Sánchez-Betancourt, L., Strategic learning and trading in broker-mediated markets. arXiv preprint arXiv:2412.20847, 2024.
- Arroyo, A., Cartea, A., Moreno-Pino, F. and Zohren, S., Deep attentive survival analysis in limit order books: Estimating fill probabilities with convolutional-transformers. *Quant. Finance*, 2024, **24**(1), 35–57.
- Bagehot, W., The only game in town. *Financ. Anal. J.*, 1971, **27**(2), 12–14.
- Barzykin, A., Bergault, P. and Guéant, O., Market-making by a foreign exchange dealer. *Risk (Cutting Edge)*, 2022.
- Barzykin, A., Bergault, P. and Guéant, O., Algorithmic market making in dealer markets with hedging and market impact. *Math. Financ.*, 2023, **33**(1), 41–79.
- Bergault, P. and Sánchez-Betancourt, L., A mean field game between informed traders and a broker. *SIAM J. Financ. Math.*, 2025, **16**(2), 358–388.
- Butz, M. and Oomen, R., Internalisation by electronic FX spot dealers. *Quant. Finance*, 2019, **19**(1), 35–56.
- Cartea, Á. and Sánchez-Betancourt, L., Optimal execution with stochastic delay. *Finance Stochastics*, 2023, **27**(1), 1–47.
- Cartea, Á. and Sánchez-Betancourt, L., Brokers and informed traders: dealing with toxic flow and extracting trading signals. *SIAM J. Financ. Math.*, 2025, **16**(2), 243–270.
- Cartea, Á., Drissi, F. and Osselin, P., Bandits for algorithmic trading with signals. Available at SSRN 4484004, 2023.
- Cartea, Á., Jaimungal, S. and Sánchez-Betancourt, L., Nash equilibrium between brokers and traders. *Finance and Stochastics*, to appear, arXiv:2407.10561, 2025+.
- Copeland, T.E. and Galai, D., Information effects on the bid-ask spread. *J. Finance*, 1983, **38**(5), 1457–1469.
- Donnelly, R. and Li, Z., Liquidity competition between brokers and an informed trader. *Int. J. Theor. Appl. Fin.*, 2025, 2550018.
- Duran-Martin, G., Kara, A. and Murphy, K., Efficient online bayesian inference for neural bandits. In *International Conference on Artificial Intelligence and Statistics*, pp. 6002–6021, 2022 (PMLR: Valencia, Spain).
- Easley, D., Kiefer, N.M., O’Hara, M. and Paperman, J.B., Liquidity, information, and infrequently traded stocks. *J. Finance*, 1996, **51**(4), 1405–1436.
- Fawcett, T., An introduction to roc analysis. *Pattern Recognit. Lett.*, 2006, **27**(8), 861–874.
- Frankle, J. and Carbin, M., The lottery ticket hypothesis: Finding sparse, trainable neural networks. In *International Conference on Learning Representations*, 2019. <https://openreview.net/forum?id=rJl-b3RcF7>
- Glosten, L.R. and Milgrom, P.R., Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *J. Financ. Econ.*, 1985, **14**(1), 71–100.
- Grossman, S.J. and Stiglitz, J.E., On the impossibility of informationally efficient markets. *Am. Econ. Rev.*, 1980, **70**(3), 393–408.
- Hastie, T., Tibshirani, R. and Friedman, J., *The Elements of Statistical Learning*, Springer Series in Statistics, 2001 (Springer, New York, NY).
- Kingma, D.P. and Ba, J., Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015*, edited by Y. Bengio, and Y. LeCun, 2015 (International Conference on Learning Representations (ICLR): San Diego, CA).
- Kyle, A.S., Continuous auctions and insider trading. *Econometrica*, 1985, **53**, 1315–1335.
- Kyle, A.S., Informed speculation with imperfect competition. *Rev. Econ. Stud.*, 1989, **56**(3), 317–355.
- Lambert, M., Bonnabel, S. and Bach, F., The recursive variational Gaussian approximation (R-VGA). *Stat. Comput.*, 2021, **32**(1), 1–24.
- Larsen, B.W., Fort, S., Becker, N. and Ganguli, S., How many degrees of freedom do we need to train deep networks: A loss landscape perspective. In *International Conference on Learning Representations*, 2022. <https://openreview.net/forum?id=ChMLTGRjFcU>
- Li, C., Farkhoor, H., Liu, R. and Yosinski, J., Measuring the intrinsic dimension of objective landscapes. In *International Conference on Learning Representations*, 2018. <https://openreview.net/forum?id=ryup8-WCW>
- Lin, W., Khan, M.E. and Schmidt, M., Stein’s lemma for the reparameterization trick with exponential family mixtures. Preprint arXiv:1910.13398, 2019.
- Liu, D.C. and Nocedal, J., On the limited memory BFGS method for large scale optimization. *Math. Program.*, 1989, **45**(1–3), 503–528.
- McElfresh, D., Khandagale, S., Valverde, J., Prasad C, V., Ramakrishnan, G., Goldblum, M. and White, C., When do neural nets outperform boosted trees on tabular data? *Adv. Neural Inf. Process. Syst.*, 2023, **36**, 76336–76369.
- Murphy, K.P., *Probabilistic Machine Learning: An Introduction*, 2022 (MIT Press: Cambridge, MA).
- Murphy, K.P., *Probabilistic Machine Learning: Advanced Topics*, 2023 (MIT Press: Cambridge, MA).
- Ollivier, Y., Online natural gradient as a Kalman filter. *Electron. J. Stat.*, 2018, **12**, 2930–2961.
- Oomen, R., Execution in an aggregator. *Quant. Finance*, 2017, **17**(3), 383–404.
- West, R.M., Best practice in statistics: The use of log transformation. *Ann. Clin. Biochem.*, 2022, **59**(3), 162–165.
- Wu, X. and Jaimungal, S., Broker-trader partial information nash equilibria. arXiv preprint arXiv:2412.17712, 2024.

## Appendices

## Appendix 1. Features

For a given trading day  $\vartheta \in \mathfrak{D}$ , the processes

$$\left(S_t^{a,\vartheta}\right)_{t \in \mathfrak{T}}, \quad \left(S_t^{b,\vartheta}\right)_{t \in \mathfrak{T}}, \quad \left(V_t^{a,\vartheta}\right)_{t \in \mathfrak{T}}, \quad \left(V_t^{b,\vartheta}\right)_{t \in \mathfrak{T}},$$

denote the best ask price, the best bid price, the volume at the best ask price, and the volume at the best bid price in LMAX Exchange for day  $\vartheta$ , respectively — we drop the superscript  $\vartheta$  when we do not wish to draw attention to the day. The feature associated with the log-transformed inventory of client  $c \in \mathcal{A}$  is

$$\text{sign}(Q_{t-}^c) \times \log(1 + |Q_{t-}^c|),$$

where  $Q_{t-}^c$  is the position in lots (one lot is  $\in 10,000$ ) of client  $c$  accumulated over  $[0, t)$ . Here,  $q_t^c$  is the size of the order sent at time  $t$  by client  $c$  and  $N_t^{c,a}$ ,  $N_t^{c,b}$  are the counting processes of buy and sell orders, respectively, sent by client  $c$  and filled by the broker. The cash of client  $c$ , denoted by  $C_t^c$ , is given by

$$C_t^c = - \int_0^t S_{u-}^a q_u^c dN_u^{a,c} + \int_0^t S_{u-}^b q_u^c dN_u^{b,c},$$

and the feature associated with the cash process is

$$\text{sign}(C_{t-}^c) \times \log(1 + |C_{t-}^c|).$$

In LMAX Exchange, the bid-ask spread is

$$S_t = S_t^a - S_t^b,$$

the midprice is

$$S_t = \frac{S_t^a + S_t^b}{2},$$

the volume imbalance of the best available volumes is defined by

$$\mathcal{I}_t = \frac{V_t^b - V_t^a}{V_t^b + V_t^a},$$

and the associated feature for the volume  $V$  is the transformed volume

$$\mathcal{V} = \log(1 + |V|).$$

The number of trades received by the broker from her clients is

$$N_t = \sum_{c \in \mathcal{A}} \left( N_t^{c,a} + N_t^{c,b} \right),$$

and the volatility of the midprice in the LOB of LMAX Exchange over the interval  $[t - \delta, t]$  is the square root of the quadratic variation of the logarithm of the midprice over the interval. More precisely,

$$\mathcal{V}_t^\delta = \sqrt{\sum_{\Delta \log S_u \neq 0; u \in [t-\delta, t]} |\Delta \log S_u|^2},$$

where

$$\Delta \log S_u = \log S_u - \log S_{u-}, \quad \text{and} \quad S_{u-} = \lim_{v \nearrow u} S_v.$$

The return of the exchange rate of the currency pair over a period  $\delta > 0$  is given by

$$\mathcal{R}_t^\delta = \log(S_t / S_{t-\delta}).$$

The timing of the events in the LOB is measured with three clocks: time-clock, transaction-clock, and volume-clock. The time-clock runs as  $t \in [0, T]$  with microsecond precision, i.e. a millionth of a second. For a given day  $d$  with  $N^d$  transactions and  $V^d$  volume traded, the transaction clock runs as  $t \in [0, N^d]$ , and the volume-clock runs as  $v \in [0, V^d]$ . The number of transactions  $\mathfrak{t}(t)$  is the number of transactions observed up until time  $t$ , similar for the volume-clock  $\mathfrak{v}(t)$ . Thus, for any order sent at time  $t$ , the time associated with the order in the transaction-clock is  $\mathfrak{t}(t)$  and the time in the volume-clock is  $\mathfrak{v}(t)$ .

For each clock  $c \in \{\text{transaction, time, volume}\}$ , we build features spanning an interval  $[\ell_c 2^n, \ell_c 2^{n+1})$  of units in the respective clock with  $\ell_c > 0$ , and use a given statistic to summarise the values in the interval; for example, for spread, imbalance, and transformed volumes we use the average value over the period. In our experiments, we consider seven intervals that span the ranges  $[0, \ell_c)$  and  $\{[\ell_c 2^n, \ell_c 2^{n+1})\}_{n=0}^6$ . The median time elapsed between any two transactions for the six clients is 1.8 seconds and the median quantity traded with LMAX Broker is €2,000. Thus,  $\ell_{\text{transaction}}$  is one transaction,  $\ell_{\text{time}}$  is one second, and  $\ell_{\text{volume}}$  is €2,000.†

## Appendix 2. PULSE derivation and proofs

In this section, we present the derivation and proofs for PULSE. Proposition A.1 derives the general fixed point equations. Given

† The intervals for the transaction-clock are  $[0, 1)$  transaction,  $[1, 2)$  transactions,  $[2, 4)$  transactions,  $\dots$ , and  $[2^5, 2^6)$  transactions. Similarly, the intervals for the time-clock are  $[0, 1)$  second,  $[1, 2)$  seconds,  $[2, 4)$  seconds,  $\dots$ , and  $[2^5, 2^6)$  seconds. Lastly, the intervals in the volume-clock are  $[0, e2000)$ ,  $[e2000, e4000)$ ,  $[e4000, e8000)$ ,  $\dots$ , and  $[e2000 \times 2^5, e2000 \times 2^6)$ .

that these are expensive to compute, we present additional results to obtain the computationally efficient form of the theorem. We then prove Theorem 2 in A.1.

**PROPOSITION A.1 (Modified R-VGA for PULSE)** *Suppose  $\log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)$  is differentiable with respect to  $(\mathbf{z}, \mathbf{w})$  and the observations  $\{y_n\}_{n=1}^N$  are conditionally independent over  $(\mathbf{z}, \mathbf{w})$ . Given Gaussian prior distributions  $\phi_0, \varphi_0$  for  $\mathbf{w}$  and  $\mathbf{z}$  respectively, the variational posterior distributions at time  $n \in \{1, \dots, N\}$  that solve (9) satisfy the fixed-point equations*

$$\begin{aligned} \mathbf{v}_n &= \mathbf{v}_{n-1} - \Sigma_{n-1} \nabla_{\mathbf{v}} \mathbb{E}_{\phi_n(\mathbf{w}), \varphi_n(\mathbf{z})} [\log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)], \\ \boldsymbol{\mu}_n &= \boldsymbol{\mu}_{n-1} - \Gamma_{n-1} \nabla_{\boldsymbol{\mu}} \mathbb{E}_{\phi_n(\mathbf{z}), \varphi_n(\mathbf{z})} [\log p(y_n | \mathbf{z}, \boldsymbol{\psi}; \mathbf{x}_n)], \\ \Sigma_n^{-1} &= \Sigma_{n-1}^{-1} + 2 \nabla_{\Sigma} \mathbb{E}_{\phi_n(\mathbf{w}), \varphi_n(\mathbf{z})} [\log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)], \\ \Gamma_n^{-1} &= \Gamma_{n-1}^{-1} + 2 \nabla_{\Gamma} \mathbb{E}_{\phi_n(\mathbf{z}), \varphi_n(\mathbf{z})} [\log p(y_n | \mathbf{z}, \boldsymbol{\psi}; \mathbf{x}_n)]. \end{aligned} \quad (\text{A1})$$

*Proof* First, rewrite (9). Let  $p(y_n) \equiv p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)$  to simplify notation. The loss function is

$$\begin{aligned} \mathcal{K}_n &= \text{KL}(\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) || \phi_{n-1}(\mathbf{w}) \varphi_{n-1}(\mathbf{z}) p(y_n)) \\ &= \iint \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) \\ &\quad \times \log \left( \frac{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)}{\varphi_{n-1}(\mathbf{z}) \phi_{n-1}(\mathbf{w}) p(y_n)} \right) d\mathbf{z} d\mathbf{w} \\ &= \iint \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) \left[ \log \left( \frac{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma)}{\varphi_{n-1}(\mathbf{z})} \right) \right. \\ &\quad \left. + \log \left( \frac{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma)}{\phi_{n-1}(\mathbf{w})} \right) - \log p(y_n) \right] d\mathbf{z} d\mathbf{w} \\ &= \int \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \log \left( \frac{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma)}{\varphi_{n-1}(\mathbf{z})} \right) d\mathbf{z} \\ &\quad + \int \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) \log \left( \frac{\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)}{\phi_{n-1}(\mathbf{w})} \right) d\mathbf{w} \\ &\quad + \iint \mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) \log p(y_n) d\mathbf{w} d\mathbf{z} \\ &= \mathbb{E}_{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma)} \left[ \log \left( \frac{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma)}{\varphi_{n-1}(\mathbf{z})} \right) \right] \\ &\quad + \mathbb{E}_{\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)} \left[ \log \left( \frac{\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)}{\phi_{n-1}(\mathbf{w})} \right) \right] \\ &\quad + \mathbb{E}_{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)} [\log p(y_n)] \\ &= \text{KL}(\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) || \phi_{n-1}(\mathbf{w})) + \text{KL}(\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) || \varphi_{n-1}(\mathbf{z})) \\ &\quad + \mathbb{E}_{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)} [\log p(y_n)]. \end{aligned} \quad (\text{A2})$$

The first and second terms in (A2) correspond to a Kullback–Leibler divergence between two multivariate Gaussians. The last term corresponds to the posterior-predictive marginal log-likelihood for the  $t$ -th observation. To minimise (A2), we recall that the Kullback–Leibler divergence between two multivariate Gaussian is given by

$$\begin{aligned} &\text{KL}(\mathcal{N}(\mathbf{x} | \mathbf{m}_1, \mathbf{S}_1) || \mathcal{N}(\mathbf{x} | \mathbf{m}_2, \mathbf{S}_2)) \\ &= \frac{1}{2} \left[ \text{Tr}(\mathbf{S}_2^{-1} \mathbf{S}_1) + (\mathbf{m}_2 - \mathbf{m}_1)^\top \mathbf{S}_2^{-1} (\mathbf{m}_2 - \mathbf{m}_1) \right. \\ &\quad \left. - M + \log(|\mathbf{S}_2|/|\mathbf{S}_1|) \right], \end{aligned}$$

see Section 6.2.3 in Murphy (2022).

To simplify notation, let  $\mathbb{E}_{\mathcal{N}(\mathbf{z} | \boldsymbol{\mu}, \Gamma) \mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma)} [\log p(y_n)] =: \mathcal{E}_n$ . The derivative of  $\mathcal{K}_n$  with respect to  $\mathbf{v}$  is

$$\begin{aligned} \nabla_{\mathbf{v}} \mathcal{K}_n &= \nabla_{\mathbf{v}} (\text{KL}(\mathcal{N}(\mathbf{w} | \mathbf{v}, \Sigma) || \phi_{n-1}(\mathbf{w})) + \mathcal{E}_n) \\ &= \nabla_{\mathbf{v}} \left( \frac{1}{2} \mathbf{v}^\top \Sigma_{n-1}^{-1} \mathbf{v} - \mathbf{v}^\top \Sigma_{n-1}^{-1} \mathbf{v}_{n-1} + \nabla_{\mathbf{v}} \mathcal{E}_n \right) \\ &= \Sigma_{n-1}^{-1} \mathbf{v} - \Sigma_{n-1}^{-1} \mathbf{v}_{n-1} + \nabla_{\mathbf{v}} \mathcal{E}_n \end{aligned}$$

$$= \Sigma_{n-1}^{-1} (\mathbf{v} - \mathbf{v}_{n-1} - \Sigma_{n-1} \nabla_{\mathbf{v}} \mathcal{E}_n). \quad (\text{A3})$$

Set (A3) to zero and solve for

$$\mathbf{v} = \mathbf{v}_{n-1} - \Sigma_{n-1} \nabla_{\mathbf{v}} \mathcal{E}_n.$$

Next, we estimate the condition for  $\Sigma$ . Use (A2) to obtain

$$\begin{aligned} \nabla_{\Sigma} \mathcal{K}_n &= \nabla_{\Sigma} \left( -\frac{1}{2} \log |\Sigma| + \frac{1}{2} \text{Tr} \left( \Sigma \Sigma_{n-1}^{-1} \right) + \mathcal{E}_n \right) \\ &= -\frac{1}{2} \Sigma^{-1} + \frac{1}{2} \Sigma_{n-1}^{-1} + \nabla_{\Sigma} \mathcal{E}_n. \end{aligned} \quad (\text{A4})$$

The fixed-point solution for (A4) satisfies

$$\Sigma^{-1} = \Sigma_{n-1}^{-1} + 2 \nabla_{\Sigma} \mathcal{E}_n.$$

We derive the fixed-point conditions for  $\mu$  and  $\Gamma$  similarly.  $\blacksquare$

**COROLLARY 6** *Suppose  $\log p(y | \mathbf{z}, \mathbf{w})$  is differentiable with respect to  $(\mathbf{z}, \mathbf{w})$  and the observations  $\{y_n\}_{n=1}^T$  are conditionally independent over  $(\mathbf{z}, \mathbf{w})$ . Given Gaussian prior distributions  $\phi_0, \varphi_0$  for  $\mathbf{w}$  and  $\mathbf{z}$  respectively, the modified R-VGA equations for PULSE in terms of gradients and Hessians with respect to  $\mathbf{z}$  and  $\mathbf{w}$  are*

$$\begin{aligned} \mathbf{v}_n &= \mathbf{v}_{n-1} + \Sigma_{n-1} \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_n(\mathbf{z})} [\nabla_{\mathbf{w}} \log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)], \\ \mu_n &= \mu_{n-1} + \Gamma_{n-1} \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_n(\mathbf{z})} [\nabla_{\mathbf{z}} \log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)], \\ \Sigma_n^{-1} &= \Sigma_{n-1}^{-1} + \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_n(\mathbf{z})} [\nabla_{\mathbf{w}}^2 \log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)], \\ \Gamma_n^{-1} &= \Gamma_{n-1}^{-1} + \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_n(\mathbf{z})} [\nabla_{\mathbf{z}}^2 \log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)]. \end{aligned}$$

*Proof* The proof follows directly by rearranging the order of integration, Bonnet's Theorem, and Prices's Theorem. see Theorems 3 and 4 in Lin *et al.* (2019).  $\blacksquare$

Corollary A.1 provides tractable fixed-point equations. Note that in Proposition A.1 the gradients are taken with respect to model parameters outside the expectation, whereas in Corollary A.1 the gradients and Hessians are taken inside the expectation.

**PROPOSITION A.2** *Given a logistic regression model written in the canonical form*

$$\log p(y_n) = y_n \log \left( \frac{\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)}{1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)} \right) - (1 + \exp(\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n))),$$

*we have that the gradient of  $\log \left( \frac{\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)}{1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)} \right)$  with respect to  $\boldsymbol{\theta}$  is given by*

$$\nabla_{\boldsymbol{\theta}} \log \left( \frac{\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)}{1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)} \right) = \mathbf{x}_n.$$

*Proof* Use the identity  $\frac{d}{dx} \sigma(x) = \sigma(x)(1 - \sigma(x)) =: \sigma'(x)$  and write

$$\begin{aligned} \nabla_{\boldsymbol{\theta}} \log \left( \frac{\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)}{1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)} \right) &= \nabla_{\boldsymbol{\theta}} \log(\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)) - \nabla_{\boldsymbol{\theta}} \log(1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)) \\ &= (\sigma(\boldsymbol{\theta}^\top \mathbf{x}_n))^{-1} \sigma'(\boldsymbol{\theta}^\top \mathbf{x}_n) \mathbf{x}_n + (1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n))^{-1} \sigma'(\boldsymbol{\theta}^\top \mathbf{x}_n) \mathbf{x}_n \\ &= (1 - \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n) + \sigma(\boldsymbol{\theta}^\top \mathbf{x}_n)) \mathbf{x}_n \\ &= \mathbf{x}_n. \end{aligned}$$

### A.1. Proof of Theorem 2

*Proof* First rewrite the log-likelihood of the target variable  $y_n$  as a member of the exponential-family. Let  $f_n(\mathbf{z}, \mathbf{w}) = \mathbf{w}^\top h(\mathbf{z}; \mathbf{x}_n)$ . Then,

$$\log p(y_n) = \log \text{Bern}(y_n | \sigma(\mathbf{w}^\top h(\mathbf{z}; \mathbf{x}_n)))$$

$$\begin{aligned} &= y_n \log \sigma(f_n(\mathbf{z}, \mathbf{w})) + (1 - y_n) \log(1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))) \\ &= y_n \log \left( \frac{\sigma(f_n(\mathbf{z}, \mathbf{w}))}{1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))} \right) + \log(1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))) \\ &= y_n \log \left( \frac{\sigma(f_n(\mathbf{z}, \mathbf{w}))}{1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))} \right) \\ &\quad - \log \left( 1 + \exp \left( \log \left( \frac{\sigma(f_n(\mathbf{z}, \mathbf{w}))}{1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))} \right) \right) \right) \\ &= y_n \eta_n - \log(1 + \exp(\eta_n)) \\ &= y_n \eta_n - A(\eta_n), \end{aligned}$$

where  $\eta_n = \log \left( \frac{\sigma(f_n(\mathbf{z}, \mathbf{w}))}{1 - \sigma(f_n(\mathbf{z}, \mathbf{w}))} \right)$  is the natural parameter and  $A(\eta_n)$  the log-partition function. We perform a moment-match estimation. To do this, we follow the property of exponential-family distributions, and use the results in Proposition A.2. The first and second-order derivatives of the log-partition  $A(\eta_n)$  are

$$\frac{\partial}{\partial \eta_n} A(\eta_n) = \mathbb{E}[y | \eta_n] = \sigma(f_n(\mathbf{z}, \mathbf{w})), \quad (\text{A5})$$

$$\frac{\partial^2}{\partial \eta_n^2} A(\eta_n) = \text{Cov}[y | \eta_n] = \sigma'(f_n(\mathbf{z}, \mathbf{w})). \quad (\text{A6})$$

Then, the first order approximation  $\hat{\sigma}(f_n(\mathbf{z}, \mathbf{w}))$  is

$$\begin{aligned} \hat{\sigma}(f_n(\mathbf{z}, \mathbf{w})) &= \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1}) (F_{n,\mathbf{z}}^\top (\mathbf{z} - \mu_{n-1}) \\ &\quad + F_{n,\mathbf{w}}^\top (\mathbf{w} - \mathbf{v}_{n-1})), \end{aligned}$$

where  $\bar{f}_{n-1} = f_n(\mu_{n-1}, \mathbf{v}_{n-1})$ ,  $F_{n,\mathbf{z}} = \nabla_{\mathbf{z}} f_n(\mathbf{z}, \mathbf{v}_{n-1})|_{\mathbf{z}=\mu_{n-1}}$ , and  $F_{n,\mathbf{w}} = \nabla_{\mathbf{w}} f_n(\mu_{n-1}, \mathbf{w})|_{\mathbf{w}=\mathbf{v}_{n-1}} = h(\mu_{n-1}; \mathbf{x}_n)$ . Next, we derive the update equations for  $\mathbf{v}$  and  $\Sigma$ . The moment-matched log-likelihood is given by

$$\log p(y_n) = \log \mathcal{N}(y_n | \hat{\sigma}(f_n(\mathbf{z}, \mathbf{w})), \sigma(f_n(\mathbf{z}, \mathbf{w}))(1 - \sigma(f_n(\mathbf{z}, \mathbf{w})))) \quad (\text{A7})$$

and the gradient with respect to  $\mathbf{w}$  is

$$\nabla_{\mathbf{w}} \log p(y_n) = y_n F_{n,\mathbf{w}} - \hat{\sigma}(\bar{f}_{n-1}) F_{n,\mathbf{w}} = (y_n - \hat{\sigma}(\bar{f}_{n-1})) F_{n,\mathbf{w}}.$$

Now, the Hessian of the log-model with respect to  $\mathbf{w}$  is

$$\begin{aligned} \nabla_{\mathbf{w}}^2 \log p(y_n) &= \nabla_{\mathbf{w}} (y_n - \hat{\sigma}(\bar{f}_{n-1})) F_{n,\mathbf{w}} \\ &= -\sigma'(\bar{f}_{n-1}) F_{n,\mathbf{w}}^\top F_{n,\mathbf{w}}. \end{aligned}$$

The update for  $\Sigma_n$ , following the order-2 form of the modified equations and replacing the expectation under  $\varphi_n(\mathbf{z})$  for  $\varphi_{n-1}(\mathbf{z})$ , is

$$\begin{aligned} \Sigma_n^{-1} &= \Sigma_{n-1}^{-1} - \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_{n-1}(\mathbf{z})} \left[ \nabla_{\mathbf{w}}^2 \log p(y_n) \right] \\ &= \Sigma_{n-1}^{-1} - \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_{n-1}(\mathbf{z})} \left[ -\sigma'(\bar{f}_{n-1}) F_{n,\mathbf{w}}^\top F_{n,\mathbf{w}} \right] \\ &= \Sigma_{n-1}^{-1} + \sigma'(\bar{f}_{n-1}) F_{n,\mathbf{w}}^\top F_{n,\mathbf{w}} \\ &= \Sigma_{n-1}^{-1} + \sigma'(\mathbf{v}_{n-1}^\top h(\mu_{n-1}; \mathbf{x}_n)) h(\mu_{n-1}; \mathbf{x}_n)^\top h(\mu_{n-1}; \mathbf{x}_n). \end{aligned}$$

Next, the update step for  $\mathbf{v}_n$  becomes

$$\begin{aligned} \mathbf{v}_n &= \mathbf{v}_{n-1} - \Sigma_{n-1} \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_{n-1}(\mathbf{z})} [\nabla_{\mathbf{w}} \log p(y_n | \mathbf{z}, \mathbf{w}; \mathbf{x}_n)] \\ &= \mathbf{v}_{n-1} + \Sigma_{n-1} \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_{n-1}(\mathbf{z})} [(y_n - \hat{\sigma}(\bar{f}_{n-1})) F_{n,\mathbf{z}}] \\ &= \mathbf{v}_{n-1} + \Sigma_{n-1} \mathbb{E}_{\phi_n(\mathbf{w}) \varphi_{n-1}(\mathbf{z})} [(y_n - \sigma(\bar{f}_{n-1})) \\ &\quad + \sigma'(\bar{f}_{n-1}) (F_{n,\mathbf{z}} (\mathbf{z} - \mu_{n-1}) + F_{n,\mathbf{w}}^\top (\mathbf{w} - \mathbf{v}_{n-1}))] F_{n,\mathbf{z}} \\ &= \mathbf{v}_{n-1} + \Sigma_{n-1} F_{n,\mathbf{w}} (y_n - \sigma(\bar{f}_{n-1}) - \sigma'(\bar{f}_{n-1}) F_{n,\mathbf{w}}^\top (\mathbf{v}_n - \mathbf{v}_{n-1})) \\ &= \mathbf{v}_{n-1} + \Sigma_{n-1} F_{n,\mathbf{w}} (y_n - \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1}) F_{n,\mathbf{w}}^\top \mathbf{v}_{n-1}) \\ &\quad - \sigma'(\bar{f}_{n-1}) \Sigma_{n-1} F_{n,\mathbf{w}} F_{n,\mathbf{w}}^\top \mathbf{v}_n. \end{aligned}$$

We rewrite the last equality as

$$\begin{aligned} \mathbf{v}_n + \sigma'(\bar{f}_{n-1})\boldsymbol{\Sigma}_{n-1}F_{n,w}F_{n,w}^T\mathbf{v}_n \\ = \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1}), \end{aligned}$$

which implies that

$$\begin{aligned} (\mathbf{I} + \sigma'(\bar{f}_{n-1})\boldsymbol{\Sigma}_{n-1}F_{n,w}F_{n,w}^T)\mathbf{v}_n \\ = \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1}). \end{aligned}$$

Similarly, we have that

$$\begin{aligned} \mathbf{v}_n = (\mathbf{I} + \sigma'(\bar{f}_{n-1})\boldsymbol{\Sigma}_{n-1}F_{n,w}F_{n,w}^T)^{-1} \\ (\mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1})), \end{aligned}$$

where

$$\begin{aligned} (\mathbf{I} + \sigma'(\bar{f}_{n-1})\boldsymbol{\Sigma}_{n-1}F_{n,w}F_{n,w}^T)^{-1} \\ = \left(\boldsymbol{\Sigma}_{n-1} \left[ \boldsymbol{\Sigma}_{n-1}^{-1} + \sigma'(\bar{f}_{n-1})F_{n,w}F_{n,w}^T \right]\right)^{-1} \\ = \left[ \boldsymbol{\Sigma}_{n-1}^{-1} + \sigma'(\bar{f}_{n-1})F_{n,w}F_{n,w}^T \right]^{-1} \boldsymbol{\Sigma}_{n-1}^{-1} \\ = \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_{n-1}^{-1}. \end{aligned}$$

Then, it follows that

$$\begin{aligned} \mathbf{v}_n &= \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_{n-1}^{-1} (\mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1}) \\ &\quad + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1})) \\ &= \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_{n-1}^{-1} \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_{n-1}^{-1} \boldsymbol{\Sigma}_{n-1}F_{n,w} \\ &\quad (y_n - \sigma(\bar{f}_{n-1}) + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1}) \\ &= \boldsymbol{\Sigma}_n \left( \boldsymbol{\Sigma}_{n-1}^{-1} + \sigma'(\bar{f}_{n-1})F_{n,w}^T\mathbf{v}_{n-1} \right) \mathbf{v}_{n-1} \\ &\quad + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1})) \\ &= \boldsymbol{\Sigma}_n \boldsymbol{\Sigma}_{n-1}^{-1} \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1})) \\ &= \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}F_{n,w}(y_n - \sigma(\bar{f}_{n-1})) \\ &= \mathbf{v}_{n-1} + \boldsymbol{\Sigma}_{n-1}h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n)(y_n - \sigma(\mathbf{v}_{n-1}^T h(\boldsymbol{\mu}_{n-1}; \mathbf{x}_n))). \end{aligned}$$

The updates for  $\boldsymbol{\mu}_n$  and  $\boldsymbol{\Gamma}_n$  are obtained similarly.  $\blacksquare$

### Appendix 3. Robustness analysis

#### A.2. Inventory aversion

Here, we study how the value of inventory aversion parameter  $\Phi$  affects the the AUC and the internalisation-externalisation strategy. Intuitively, as the value of inventory aversion parameter increases, the broker is willing to increase the cutoff probability to internalise a trade if it would reduce the absolute value of her inventory; similarly, she is willing to decrease the cutoff probability to internalise a trade if internalising the trade increases the absolute value of her inventory.

A higher value of the inventory aversion parameter does not necessarily mean that the broker externalises more trades, it means that the broker is keener to externalise trades that decrease the absolute value of her inventory and less keen to internalise trades that increase the absolute value of her inventory. Figure A1 (left) shows the internalised volume as a function of the inventory aversion parameter  $\Phi$ . Figure A1 (right) shows the PnL and avoided profits of the internalisation-externalisation strategy in (15) as a function of the inventory aversion parameter  $\Phi$  for a toxicity horizon of 60s using PULSE.

As the value of the inventory aversion parameter  $\Phi$  increases, the broker tends to internalise fewer trades. The performance of the

Table A1. AUC for toxicity horizon of 30s by model and by clock. EUR/USD currency pair over the period 1 August 2022 to 21 October 2022. Here, *time* is the time-clock, *txn* is the transaction-clock, and *vol* is the volume clock.

	all	time	txn	vol
PULSE	62.5	62.6	62.6	62.6
LogR	50.0	50.0	50.0	50.0
RF	56.8	56.3	53.8	53.1

internalisation-externalisation strategy is stable for small values of  $\Phi$ , but when  $\Phi$  becomes large the strategy performs poorly. To see this, observe that when  $\Phi > 0$ , a positive (negative) inventory  $Q > 0$  ( $Q < 0$ ) makes the effective cutoff probability higher for trades that increase (decrease) the inventory further; this makes the broker less likely to internalise trades in these respective directions. In our experiments, this asymmetry in the effective cutoff probability undermines the predictions made by the models which lowers the PnL. The skewing of effective cutoff probability also implies that fewer trades are internalised (see left panel). Fewer internalised trades also reduces realised PnL (on average) because the PnL of the broker is positive for the baseline experiments. At the same time, avoided profit increases because more of these profitable (for the broker) trades are externalised. In the high-inventory aversion regime the strategy is dominated by the urgency to unwind inventory, and the predictive power of the models matters less.

Figure A2 shows the daily volume of the internalisation-externalisation strategy in (15) as a function of the inventory aversion parameter  $\Phi$  for a toxicity horizon of 60s using PULSE. This is a detailed version of what the left panel of figure A1 describes. Indeed, the higher the value of the inventory aversion parameter, the less the broker internalises.

Finally, figure A3 shows the percentage of internalised volume as a function of the cutoff probability and the toxicity horizon.

#### A.3. Performance for independent clocks

In Section 3.2 we constructed 168 of the 175 features with three clocks: transaction-clock, time-clock, and volume-clock. Here, we explore the AUC of our models when we use only one of the three clocks to construct the features, we omit MLE because its predictions are feature-free. Recall that 168 out of the 183 features used in the calculations are measurements of 8 variables with three clocks and seven time horizons ( $168 = 3 \times 8 \times 7$ , see Section 3.2). Here, instead, we evaluate the model with 56 clock-features instead of 168, that is, we use either (i) the transaction-clock, (ii) the time-clock, or (iii) the volume-clock. Table A1 shows the AUCs, where we observe that the models are robust to the choice of clock.

We conclude that for PULSE, considering all three clocks does not add value. For RF on the other hand, the extra feature-engineering exercise does add value.

#### A.4. The added value of one model per client

Here, we study how model predictions change when we fit a model per client as opposed to one model for all clients. Figure A4 shows the outperformance in AUC from an individual model over the universal model (one model for all clients).

For PULSE, with the exception of Client 1, the performance of one model per client is significantly lower than that of the universal model. Indeed, the additional data are more valuable than a model per client. The universal model employs features that are built using the identity of the client, e.g. inventory and cash of client. Thus, it is more advantageous to have one model for all clients and benefit from more data points than one model per client at the expense of having fewer data points to train the individual models. This is the

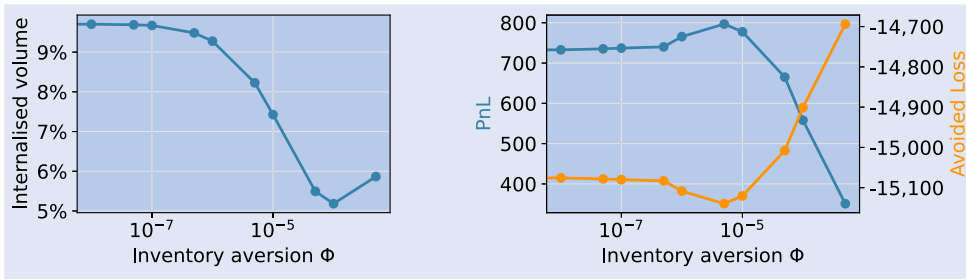


Figure A1. Left panel: percentage of internalised volume as a function of the inventory aversion parameter  $\Phi$ . Right panel: PnL and avoided profit of the internalisation-externalisation strategy as a function of the inventory aversion parameter  $\Phi$ .

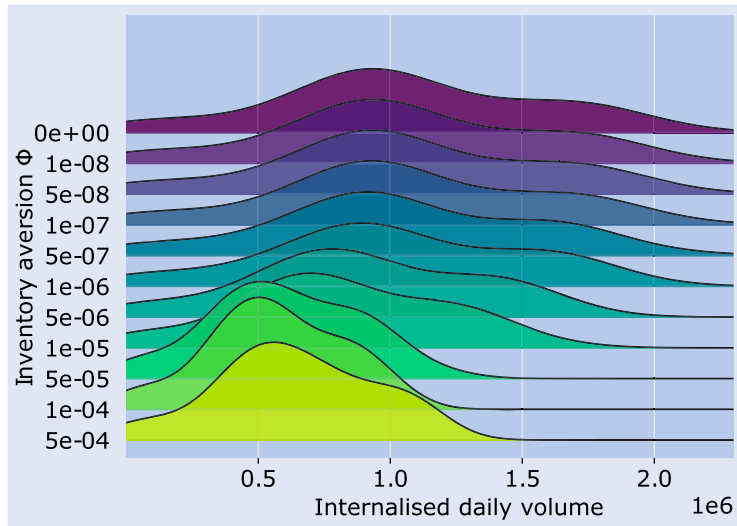


Figure A2. Kernel density estimate plot of daily traded volume for a range of values of the inventory aversion parameter  $\Phi$ .

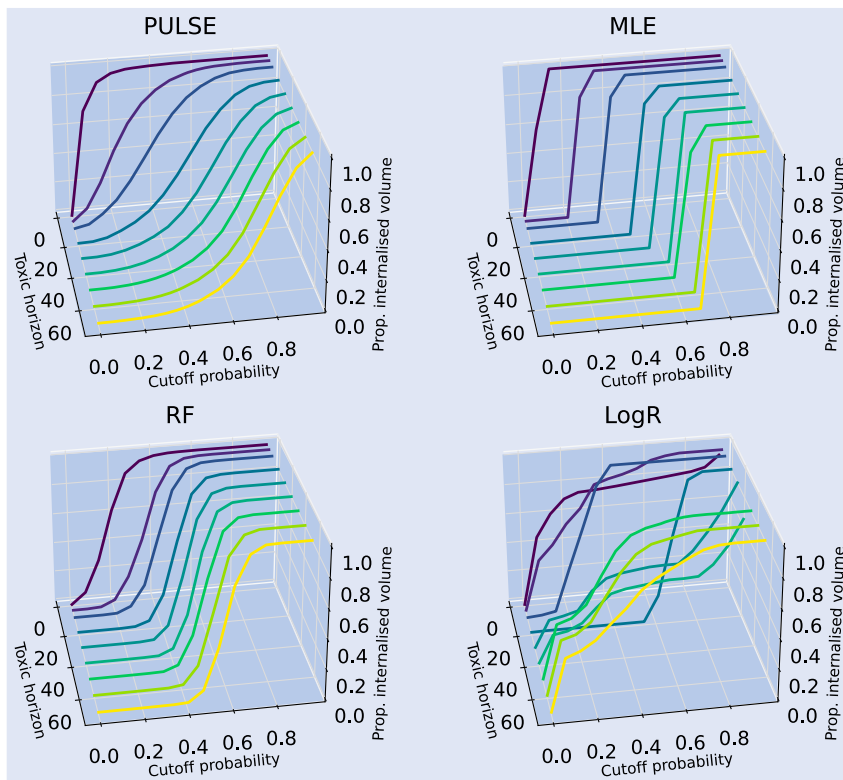


Figure A3. Proportion of internalised volume as a function of the cutoff probability  $p$  and toxicity horizon. EUR/USD currency pair over the period 1 August 2022 to 21 October 2022.

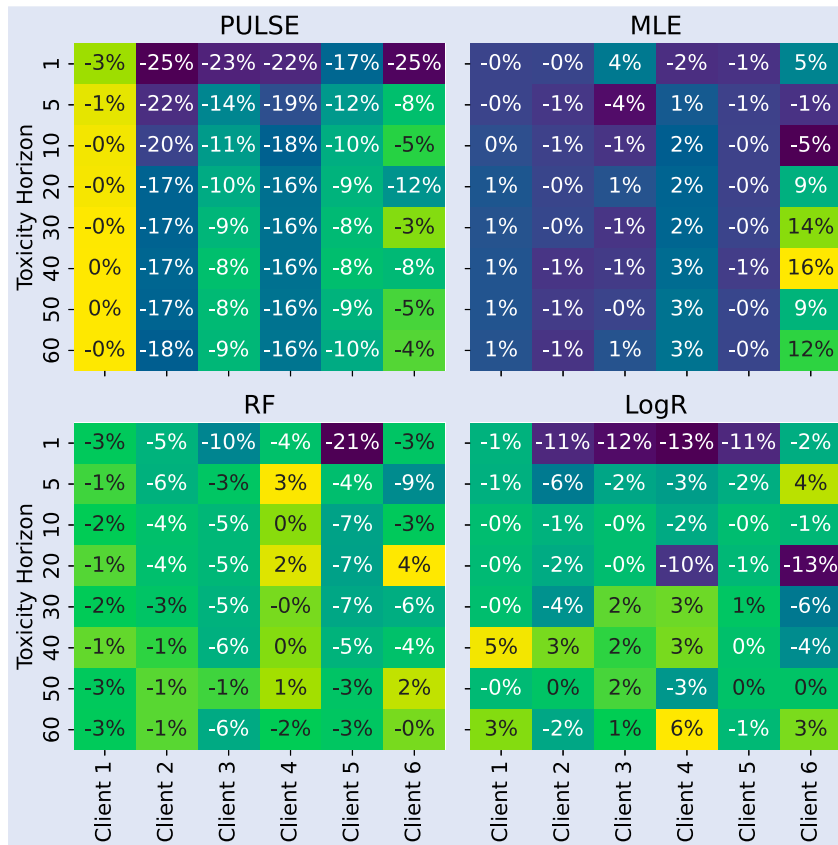


Figure A4. Ability to predict toxic flow. Difference between AUC measure without and with trader identification.

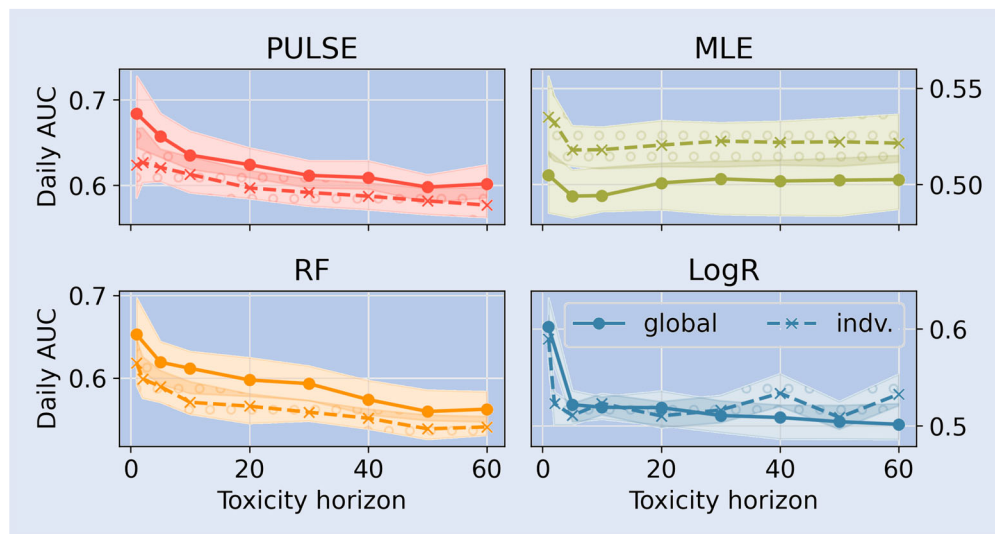


Figure A5. Each panel shows daily AUC versus toxicity horizon for one method under two configurations: (i) per-client models and (ii) a single global model. Solid lines denote the global configuration; dashed lines denote per-client; the bands around each line indicate the interquartile range (IQR). EUR/USD, 1 Aug 2022–21 Oct 2022.

case for PULSE and RF, whereas the results for LogR and MLE are not conclusive.

The above results are for a universal model that does employ client-specific features. If we consider a universal model without access to client-specific features (e.g. cash, inventory, and recent activity) the performance of the models deteriorates. For example, for Client 1 (using PULSE) we find close to 20% decrease in the ability to predict toxic flow when going from a universal model with client-specific features to a universal model without client-specific features.

### A.5. Global models

In this section we study the added value of employing data from more clients to build the models. Figure A5 reports the median daily AUC across toxicity horizons under two training regimes: (i) a single *global* model and (ii) *client-specific* models (one model per client). The global model is fit on pooled data from the top 100 clients and excludes client-identifying features. The client-specific models are trained separately for the top six clients and include client information features.

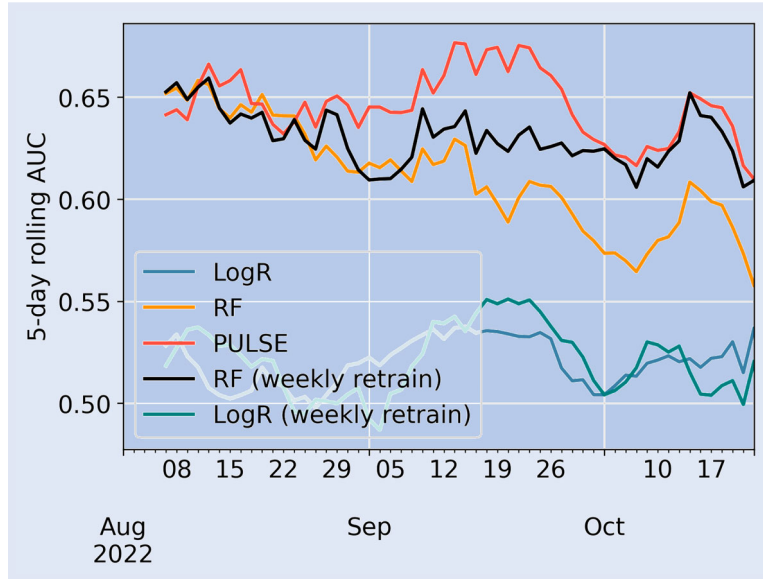


Figure A6. Five-day exponentially-weighted moving average of AUC over time. The toxicity horizon is ten seconds.

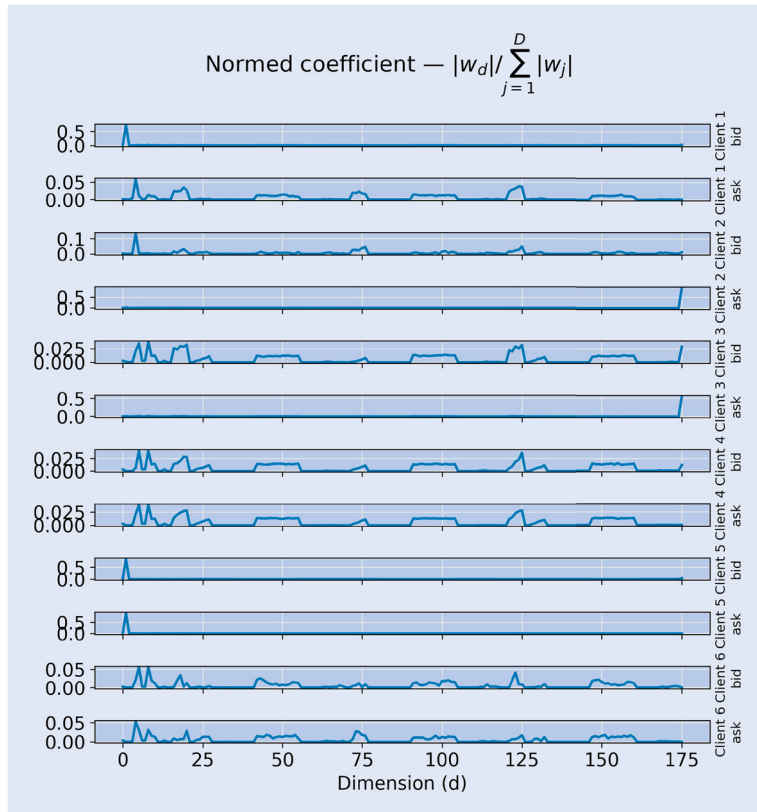


Figure A7. Normalised coefficient magnitudes  $|w_d| / \sum_j |w_j|$  for per-client LogR fits (top six clients).

Pooling data helps flexible models. The mean daily AUC is higher for PULSE and RF when using a single *global* model than when fitting one model per client—these methods can exploit the larger sample and share statistical strength across clients. LogR performs roughly on par in both settings, though the per-client variant shows higher variance due to smaller sample sizes. MLE is the exception: a per-client MLE (client-specific base rate) outperforms the global MLE, since a global average dilutes client idiosyncrasies. Overall, PULSE is the top performer, followed by RF for both training regimes. Thus, in our dataset, having more clients (more data) is advantageous. This finding is specific to our

dataset (it is not universal); this might be a consequence of the pool of clients we study, the features we consider, or the period in our analysis.

#### A.6. Weekly re-trains

In this section, we compare the performance of the NNet trained with PULSE to the performance of the RF and LogR models with weekly re-training. At the beginning of each week of the deploy phase, we use the previous week's data to re-train RF and LogR. For

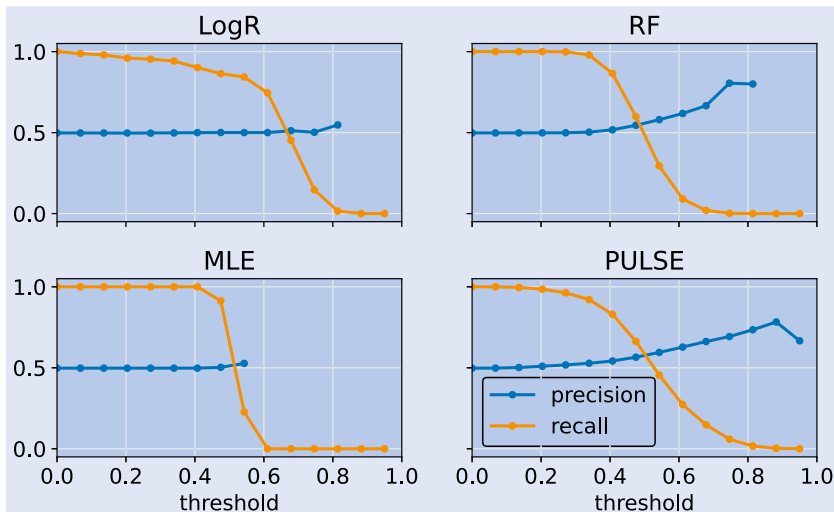


Figure A8. Precision and recall as a function of the cutoff probability at a 20s toxicity horizon, by model.

the first week in the deploy phase, we make predictions with data of the warmup phase. Figure A6 shows the five-day rolling AUC mean with  $\mathcal{G} = 10$  s. The decrease in performance for RF and LogR with weekly retrain is less pronounced than when training only in the deploy phase. As above, the mean AUC of RF and LogR with weekly re-training is lower than that of the NNet trained with PULSE.

The NNet trained with PULSE only observes each datapoint once, whereas RF keeps the data in memory to re-train. Thus, a NNet trained with PULSE is more memory-efficient and more scalable than both RF and LogR. Finally, given that PULSE is fully-online, it reacts more quickly to changes in the behaviour of clients.

### A.7. Variance of LogR

Here, we investigate the variance we observed in figure 9, A3, and the concentrated probabilities shown in figure 13.

Figure A7 shows the normalised coefficient magnitudes  $|w_j| / \sum_k |w_k|$  for per-client LogR fits. We observe a marked concentration of mass: few coefficients (in some instances, one) dominate, while the remainder are near zero, and the index of the dominant coefficient varies across clients. This reflects weak linear signal and collinearity in small per-client samples. Consistent with this, LogR’s predicted probabilities concentrate near a base rate, yielding near-step internalised-volume curves and the horizon-to-horizon jaggedness noted in the main text.

### A.8. Precision and recall

Figure A8 plots precision and recall versus the cutoff probability at a toxicity horizon of 20s. LogR and MLE exhibit flat precision and rapidly declining recall, reflecting limited score dispersion. RF and PULSE show clearer trade-offs, with precision increasing as the cut-off rises. These patterns are consistent with the internalised-volume behaviour in figure 14 and with the AUC comparisons in figure 9.