Price signatures

DEUTSCHE BANK

SALES AND TRADING DEPARTMENT

Price signatures

Roel Oomen*

September 2018

Abstract

Price signatures are statistical measurements that aim to detect systematic patterns in price dynamics localised around the point of trade execution. They are particularly useful in electronic trading because they uncover market dynamics, strategy characteristics, implicit execution costs, or counter-party trading behaviours that are often hard to identify, in part due to the vast amounts of data involved and the typically low signal to noise ratio. Because the signature summarises price dynamics over a specified time interval, it constitutes a curve (rather than a point estimate) and because of potential overlap in the price paths it has a non-trivial dependence structure which complicates statistical inference. In this paper, I show how recent advances in functional data analysis can be applied to study the properties of these signatures. To account for data dependence, I analyse and develop resampling-based bootstrap methodologies that enable reliable statistical inference and hypothesis testing. I illustrate the power of this approach using a number of case studies taken from a live trading environment in the over-the-counter currency market. I demonstrate that functional data analysis of price signatures can be used to distinguish between internalising and externalising liquidity providers in a highly effective data driven manner. This in turn can help traders to selectively engage with liquidity providers whose risk management style best aligns with their execution objectives.

*Roel Oomen is employed as a Managing Director in electronic FX spot trading at Deutsche Bank A.G. Deutsche Bank (DB) is an industry recognised world leader in the foreign exchange business, and offers a full spectrum of foreign exchange products and services, including the trading of foreign exchange products through its Autobahn electronic trading platform. This paper was prepared within the Sales and Trading function of DB, and was not produced, reviewed or edited by the DB Research Department. The views and opinions rendered in this paper reflect the author's personal views about the subject. No part of the author's compensation was, is, or will be directly related to the views expressed in this paper. See Disclaimer at end of paper. Roel Oomen would like to thank three anonymous referees, Anthony Neuberger, Andrew Patton, Jin-Ting Zhang, colleagues at Deutsche Bank, and the participants at TradeTech FX 2018 in Miami and Barcelona, the DB Quant Seminar Series, the Social and Economic Data Science seminar at the London School of Economics, the Financial Mathematics Practitioners Seminar at University College London, the Profit & Loss Forex Network 2018 in London and Chicago, Duke University, and Trinity College Cambridge for helpful comments.

1 Introduction

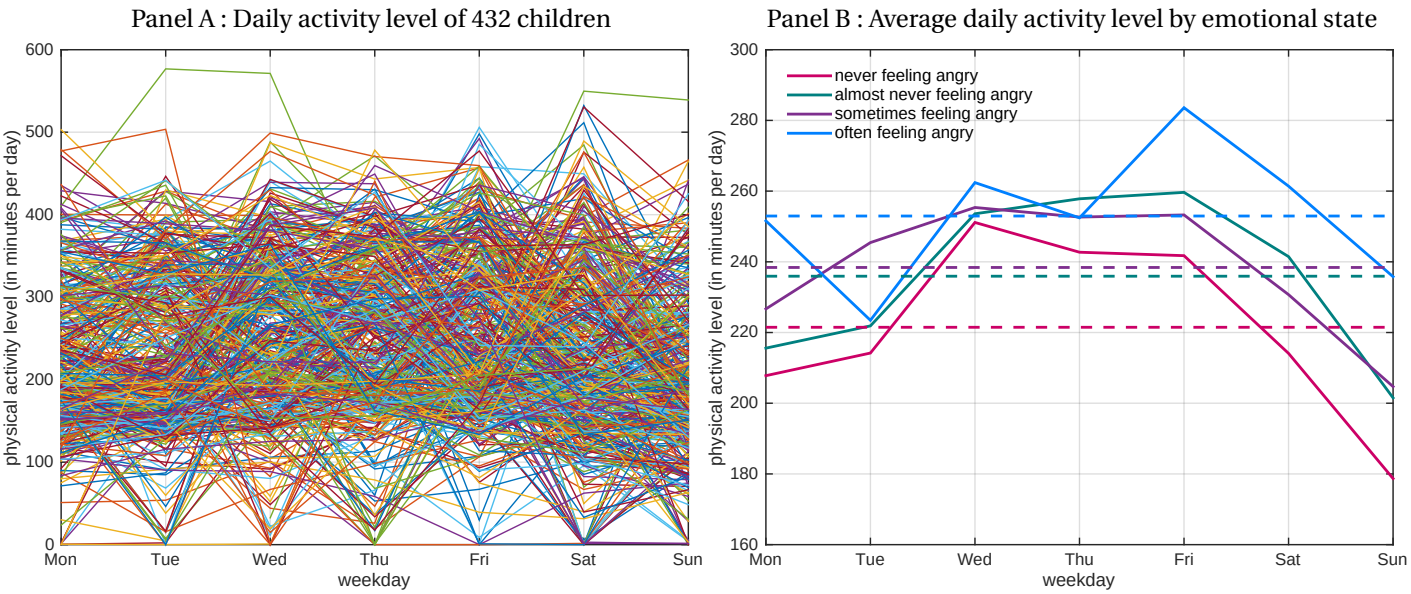
Data driven decision making is quickly turning from a best practice into a necessity in the face of a sustained growth in the amount and complexity of digital information that is being generated. Reinsel, Gantz, and Rydning (2017) estimate that in 2016 more than 15 zettabytes (ZB) of digital data was created globally, and that this figure is expected to double every three years to 44ZB by 2020 and 163ZB by 2025.[1] To put this into perspective, if one were to print out one ZB of information, it would require a stack of paper the height of more than 75 times the distance to the sun.[2] For many applications, even visual inspection of the data is no longer viable and quantitative methods are a prerequisite to search for and distill the information of interest from a vast universe. Unsurprisingly, the finance industry is subject to the same trend and this comes with a growing burden on participants to modernise workflows through an increased reliance on data driven decision making. This is perhaps most evident in electronic trading where computers that consume and exchange information at a pace approaching the speed of light have replaced humans and large transactions are completed in the blink of an eye without manual intervention. Whilst the logical design (and oversight) of such trading systems remains of course in the hands of humans, many have come to rely increasingly on quantitative data analysis to answer questions such as which venue(s) to trade on, how to optimally place orders, or how to measure true costs of execution. In this paper, I demonstrate how recent advances in the field of functional data analysis can be used to perform meaningful analysis on vast quantities of financial data to extract important and interpretable patterns that can help answer questions such as these. A number of practical case studies are used to illustrate how the signature methodology can be applied to a live trading environment. I show how the recently developed theory on aggregator execution design (Oomen, 2017a,b; Butz and Oomen, 2018) can be deployed in practice to facilitate efficient and data driven decision making that leads to measurable improvements in trading performance.

Functional data analysis, or FDA for short, provides a methodology for studying the statistical properties of data that can be represented by a curve (or a surface) as opposed to a point-wise value or a parameter. Pioneers of the contemporary literature include Ramsay and Silverman (2002, 2005). Recent textbook treatments of the topic include Horváth and Kokoszka (2012) and Zhang (2014). FDA has been successfully applied in many fields, including actuarial science, criminology, paleopathology, biology, neurology, phycology, and climatology. Figure 1 provides an illustration. Panel A draws the physical daily activity level of 432 seven year old children, as measured by an

[1] Alongside the opportunities this vast amount of data offers, it also poses a variety of challenges, including legal questions regarding ownership, privacy and security considerations, and the search for sustainable and efficient storage solutions ranging from eco-friendly data centres in the arctic circle (e.g. `http:/kolos.com`) to data encoding into bacterial DNA sequences (e.g. Extance, 2016).

[2] I assume that a page of text stores 10KB, and that a sheet of paper is 0.1mm thick. The distance to the sun is 149.6e6 km or 1 astronomical unit (AU). One ZB is thus equivalent to $n = (1024)^7/(10240)$ of pages, which when stacked would measure a height of $n/1e7/149.6e6 = 77$AU.

Figure 1: Illustration of functional data – physical activity level of seven year old children



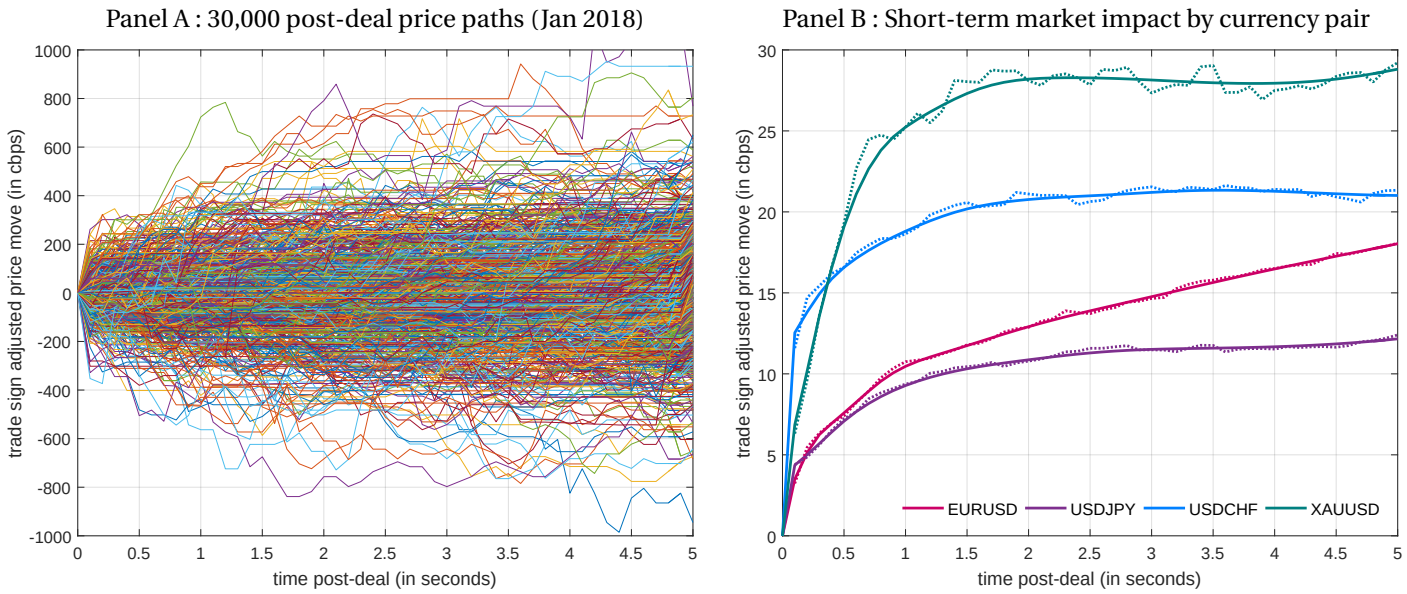
Note. Panel A draws the daily physical activity level of 432 children, measured by an accelerometer (in minutes per day) and gathered as part of the UK Millennium Cohort Study. Panel B draws the average daily activity level categorised by emotional state. The dashed lines are the weekly averages. I acknowledge: the Centre for Longitudinal Studies, Institute of Education for the use of these data; the UK Data Service for making them available; the MRC Centre of Epidemiology for Child Health (Grant reference G0400546), Institute of Child Health, University College London for creating the accelerometer data resource which was funded by the Wellcome Trust (grant reference 084686/Z/08/A). The institutions and funders acknowledged bear no responsibility for the analysis or interpretation of these data.

accelerometer worn by each child participating in the UK Millennium Cohort Study (see Griffiths et al., 2013). Panel B draws the average activity level conditioned on their emotional state. It suggests that children who often feel angry are physically more active than those who never feel angry. But are those differences statistically significant? This is an example of a question that can be addressed by FDA.[3] Figure 2 illustrates how FDA can be applied in finance. Panel A draws the trade sign adjusted price paths over the first five seconds immediately following each of 30,000 executions across a range of spot currency pairs. The sheer amount of data, combined with low signal to noise ratio, makes it impossible to visually detect any systematic directional price movement. However, after averaging the price paths as in Panel B, clear patterns appear which quantify the "market impact" of a trade, how it varies by currency pair, and how quickly it is realised post-deal. Such a curve is an example of what I define in this paper as a "price signature", and I will demonstrate that FDA is ideally suited to study its statistical properties and to conduct hypothesis testing.

The construction of a price signature is straightforward : it is computed as the volume weighted, trade direc-

[3]Appendix B provides a few more examples in Figures 14 and 15 on global warming, mortality rates, and fertility rates.

Figure 2: Illustration of functional data – market impact



Note. Panel A draws the trade sign adjusted price paths over the first five seconds immediately following each of 30,000 executions across multiple spot currency pairs. For ease of visualisation, the chart only includes a stratified subset of the full sample. Panel B draws the average signed price paths by currency pair. The dotted (solid) lines are the raw (smoothed) averages. Source : Deutsche Bank.

tion adjusted average price movement over a pre-determined interval centred around the point of trading (a formal definition of the signature is given in Eq. 1 below). Despite its simplicity, it constitutes a versatile statistic that can help to identify and characterise salient properties of a trading strategy by measuring any systematic price dynamics localised around the point of execution. I show that the signature can be used to identify and analyse a wide variety of trading strategies or market dynamics both at a macroscopic level over horizons of minutes or hours (e.g. momentum and reversal strategies, take-profit and stop-loss orders) as well as at a microscopic level of seconds or milliseconds (e.g. statistical arbitrage, adverse selection, symmetric and asymmetric last look).

Standard FDA tests can, however, only be used for signature analysis in fairly restrictive circumstances. A distinctive feature of the signature is that the signed price paths that enter into its construction are often dependent, and this violates an important IID-ness assumption required by much of the FDA theory. To understand how this dependence arises, imagine a trader who executes a hundred trades in quick succession, say one per second, and is now interested in the price signature over a one minute horizon. In this case, the price path associated with the first trade will overlap with those of the subsequent 59 trades and this is – in part – what accounts for the dependence in observations. Because the signature takes trade direction into account, any dependence structure in the sequence of buy and sell orders will also propagate into the signature. A further complicating factor is the (trade

size) weighting of observations, which is not a feature commonly encountered in the FDA literature. To investigate these effects, I specify a simple model for price and trade dynamics and derive closed form expressions for the signature variance which confirm the above intuition: the dependence structure of the signature depends on (a) the time-overlap in its constituent price paths, (b) the serial dependence in the sequence of trade signs, and (c) the variability in trade sizes. Because these non-standard data properties invalidate the commonly used FDA tests, I explore a variety of resampling-based bootstrap methods to enable statistical hypothesis testing of the signatures under weaker assumptions. Using simulations, I assess the effectiveness of the stationary block bootstrap (Politis and Romano, 1994) which is designed to account for general forms of dependence in observations. To determine the block length, I follow the procedure proposed by Politis and White (2004); Patton, Politis, and White (2009) taking as input the strength of autocorrelation in the trade sign. Next, I propose an alternative bootstrap method that exploits the fact that the source of dependence in the signature data comes from time overlap in the constituent price paths. The idea is simple: a new block is established each time there is a gap between consecutive observations that exceeds the signature horizon. This ensures that the sampled blocks are mutually uncorrelated and can then be resampled using the IID bootstrap. Because the price path overlap directly depends on the signature horizon, and thus affects the block selection, I refer to this methodology as the adaptive block bootstrap. Simulation results confirm that both the stationary bootstrap and the adaptive block bootstrap perform well in a variety of scenarios, with one outperforming the other only in terms of second order properties depending on the specific scenario.

I apply the developed FDA methodology to a number of case studies taken from a live over-the counter (OTC) trading environment. The context is as follows. In OTC financial markets, traders tend to maintain bi-lateral and disclosed relationships with multiple dealers who provide them with liquidity and compete for their flow. With advances in electronic trading, a common approach for the trader is to build an aggregator which consolidates the liquidity offered by selected dealers into a unified order book and to then use this for execution, typically on best price. Which dealers to include into the aggregator, how to execute large trades, and how to evaluate execution performance are non-trivial questions and in practice there is little consensus on how to approach these. In Oomen (2017a), I outline some of the considerations regarding aggregator design and usage from a game theoretic perspective, and conclude that as a rule of thumb, it is optimal for a trader with natural liquidity demand to limit the number of dealers in the aggregator and only to those who warehouse or internalise the risk.[4] In practice, however, it is a challenge to reliably distinguish between dealers that internalise and those that externalise. Simply asking the dealer may not be sufficient, because the concept is open to interpretation and the answer is likely to vary with

[4]Two other important recommendations are for the trader to execute each ticket for the full amount with a single selected dealer (rather than splitting it across multiple dealers simultaneously), and to leave a sufficient amount of time in between consecutive executions to allow the dealer to internalise the risk. Both are controllable by the trader and practical guidance on the latter is found in Butz and Oomen (2018).

factors that are often unobservable or unpredictable by at least one party (e.g. dealer's position at time of the trade, liquidity dynamics, properties of the trader's flow). Therefore, the approach I take is to let the data – rather than the trader or dealer – speak and extract the central message by applying FDA and associated resampling methods to the relevant signature measurements. The case studies illustrate the power of this approach. They demonstrate that it is possible to distinguish between internalisers and externalisers in a quantitative data driven manner. Moreover, such an approach can be highly effective in terms of accuracy of the conclusions reached and by the impact of the decisions it informs regarding optimal aggregator design. As predicted by the theory, all case studies result in a mutually beneficial outcome for both the trader and the dealers in that the analysis facilitates an improved alignment between the trader's execution objectives and the nature of the liquidity provided by the dealers that are selected to participate in the aggregator. In other words, it removes frictions and inefficiencies from the trading process. One important example of this is that the methods outlined in this paper enable the trader in a data-driven manner to monitor for and eliminate "prisoner's dilemma" effects in the aggregator setup (as discussed from a theoretical perspective in Oomen, 2017a).

Price signatures are only one specific example within a broader class of signatures for which the FDA methodology developed in this paper is applicable. For instance, signatures of spread, liquidity, or market activity dynamics can be constructed and studied. Similarly, signatures for rejected trades, unexecuted quotes, or trading events around macro economic announcements may be considered. As such, the results have wide applicability well beyond aggregator design, including transaction cost analysis, strategy performance measurement, the study of market integrity, stability, and liquidity conditions. Also, from a MiFID II perspective (European Parliament and the Council, 2014), the signatures can provide a quantitative characterisation of dealers classified under the systematic internaliser regime and feed into best execution considerations.

The remainder of this paper is organised as follows. Section 2 formally defines the price signature, illustrates various patterns that may be encountered in practice, analyses its theoretical properties using a simple model, and then explores how FDA can be applied to perform formal statistical testing on the signatures. Section 3 presents the case studies. Section 4 concludes. Proofs are found in Appendix A and additional tables and figures in Appendix B.

2 Price Signature

Assume a trader executes a series of transactions in a specified security at time points t_n, for a notional amount or quantity q_n, and in direction $d_n \in \{1,-1\}$ for $n = 1,\ldots,N$. Here, $d_n = 1$ when the trader buys, and $d_n = -1$ when the

trader sells. The price signature, denoted by S, associated with this set of trades is then defined as:

$$S(\delta) = \frac{1}{q'\iota} \sum_n q_n d_n (P_{t_n+\delta} - P_{t_n}), \qquad \text{for } \delta \in [-\underline{\delta}, \overline{\delta}], \tag{1}$$

where P_t is a suitably chosen logarithmic reference price (e.g. the mid) for the security at time t, $q = (q_1, q_2, \ldots, q_N)'$, and ι is a $N \times 1$ vector of ones. The parameter δ specifies the signature horizon and can be set to positive or negative values depending on whether the focus is on post-deal or pre-deal price dynamics respectively. Note that $S(0) = 0$ by construction.[5] See Figure 3 for an illustration of the signature construction. As mentioned in the introduction, in this paper I concentrate on price signatures, but the same methods can be applied to signatures of other metrics (e.g. market spread, activity, or liquidity) and calculated at non-trade events (e.g. quote placement, rejected trades, scheduled news events). As such, the methodology developed for analysis of signatures has applicability well beyond what is discussed here.

2.1 Signature interpretation & patterns

The signature in Eq. (1) is designed to identify and characterise systematic price dynamics localised around the point of trading, i.e. it is a function of both the price dynamics and the trading strategy. It can be used by a trader for a variety of purposes, for instance: to explain trading performance, identify changes in market conditions, assess strategy sensitivity to the timing of trades or system latencies, measure market impact, or to attribute revenues to individual (sets of) transactions. The signature value $S(\delta)$ can loosely be interpreted as a measure of the trader's revenue per unit size when marked-to-market at point $\delta > 0$.[6] Similarly, for $\delta < 0$, the signature measures the opportunity cost of not having executed the trades δ time-units earlier. Panel A of Figure 4 provides a graphical illustration.

The interval or horizon over which to calculate the signature is typically determined by the application at hand. For instance, if the market dynamics or strategy performance at a micro-structure level are the focus, then the signature should be calculated over a correspondingly fine-grained and short horizon, typically measured in seconds or even milli- or micro-seconds. I refer to this as a "micro signature". If, on the other hand, the analysis is focused on the macroscopic properties then the horizon can be set to minutes, hours, or even days to produce a "macro signature".

In practice, a variety of micro and macro signature patterns may be encountered. Figures 4 – 6 below provide

[5] In practice, one may incorporate the paid or earned bid-offer spread into the signature so that $S(0)$ measures the volume weighted average spread. I abstract away from this for simplicity of exposition, but all the methods presented in this paper directly apply to this case.

[6] This of course assumes that the trader holds on to their trades for longer than δ. If the trades are closed out prior to that, the incremental variation in the signature value represents the opportunity costs of not having held on to the positions longer.

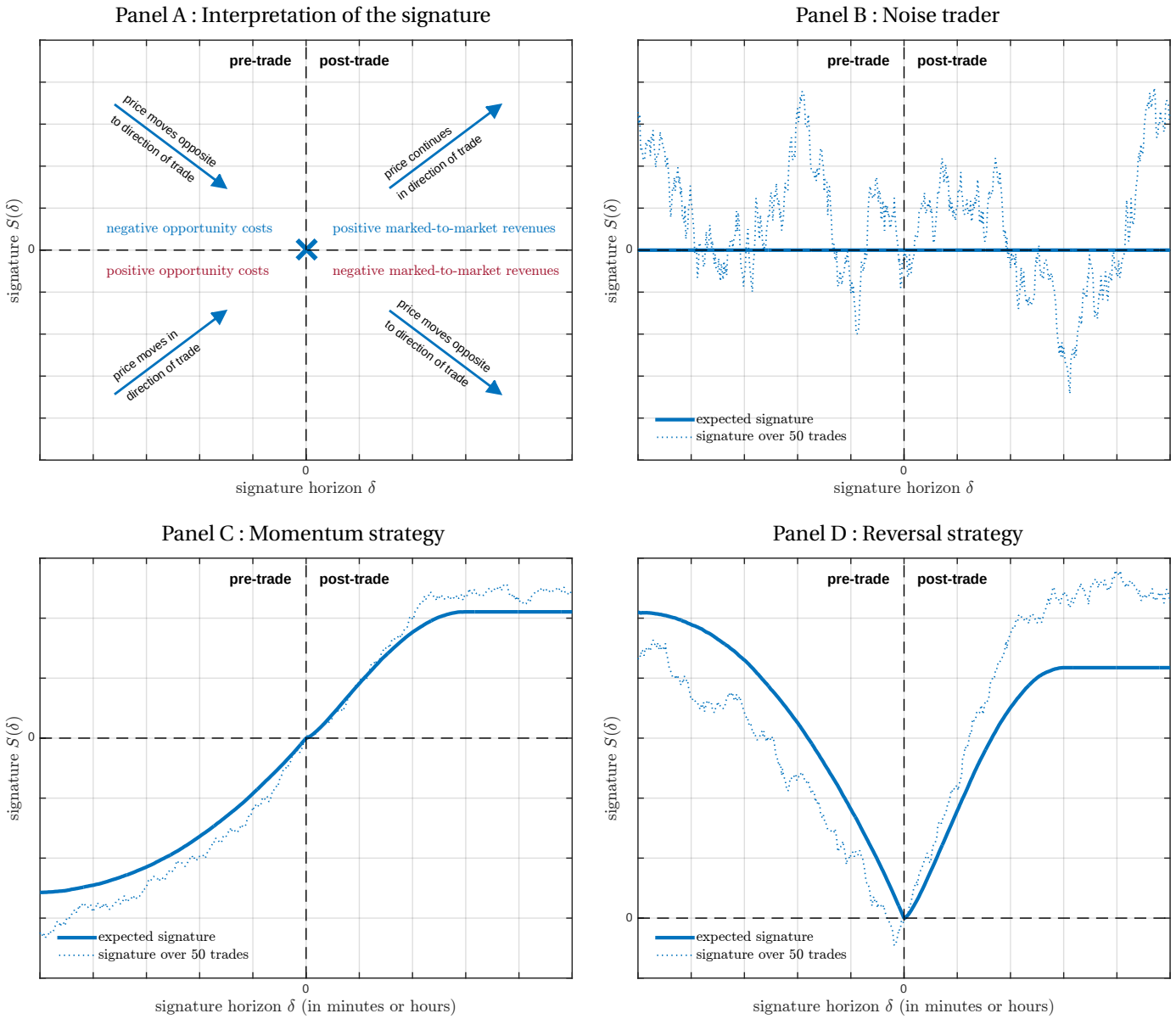
Figure 3: Illustration of price signature construction

Note. This chart draws a simulated price path together with a set of buy and sell transactions. The grey lines indicate the signature horizon (i.e. $\underline{\delta} = 0, \overline{\delta} = 60$ seconds). The shaded grey areas highlight the intervals where portions of the price paths overlap.

a number of stylised examples. All are based on simple models and simulated data. Starting with Figure 4, Panel B draws the signature associated with that of a "noise trader", i.e. the timing, direction (and size) of their trades are exogenously determined and independent of the price dynamics both before and after the trade. The expected signature is flat, although when calculated using a finite sample of data the signature will naturally include some measurement error (as illustrated by the dotted line). Panels C and D draw the signature of a momentum and reversal strategy respectively. Here, the trader determines the timing and direction of their trades based on the historic price evolution. The momentum trader expects a recent upward (downward) trend to continue and buys (sells) in anticipation. The corresponding signature is upward sloping both pre- and post-deal, reflecting in part that the later a trend is detected the higher the opportunity costs. For the reversal strategy, the trader expects local trends to revert and aims to trade at the turning points. The associated signature has a U-shaped pattern. An unsuccessful trader's signature may still exhibit a similar pre-deal pattern, but the expected post-deal signature would be flat or negative.

Rather than forming trade decisions on the recent price evolution – as in the momentum and reversal strategies – the trader may use predictive signals or "alpha" that are exogenous and uncorrelated to the pre-deal price dynamics. Panel A of Figure 5 draws an example signature associated with alpha that is realised almost immediately after trading and one that is only gradually realised. Because it is often hard to disentangle alpha from market impact

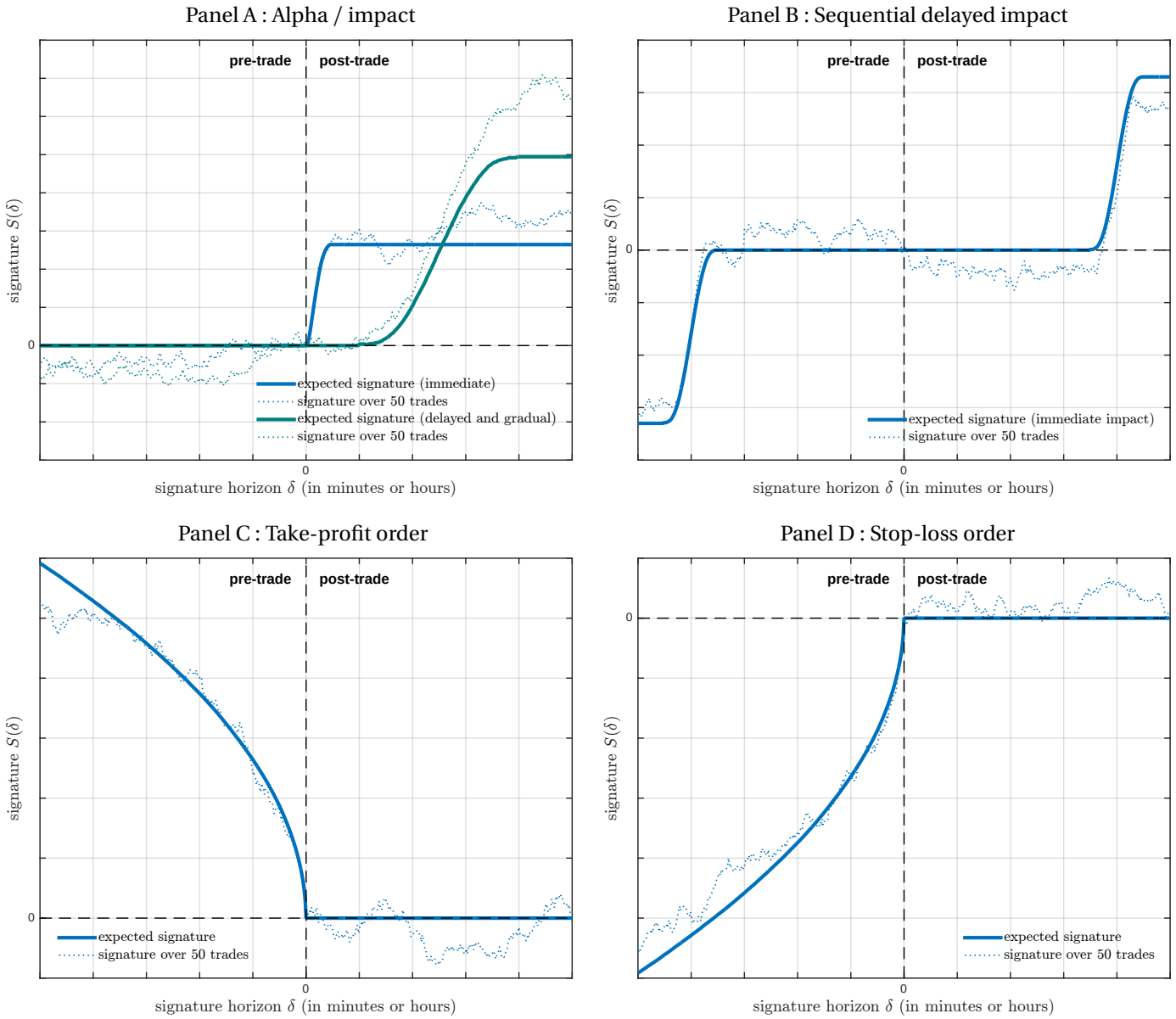
Figure 4: Signature profiles at a macroscopic level



Note. Panel A provides a qualitative interpretation of the signature profile $S(\delta)$ as defined by Eq. (1). Panels B – D provide example profiles associated with different trading strategies. The solid line is the expected signature for the assumed strategy and underlying price dynamics. The dotted line is a signature based on a set of 50 observations.

or information leakage, the same signature patterns may be observed for all such scenarios. Panel B draws a signature where the trading frequency is shorter than the signature horizon so that the post-deal impact of one trade feeds into the pre-deal signature of the subsequent trade. An example of this would be where the trader executes a sequence of block trades, starting every block execution with minimal impact and then gradually increasing the ag-

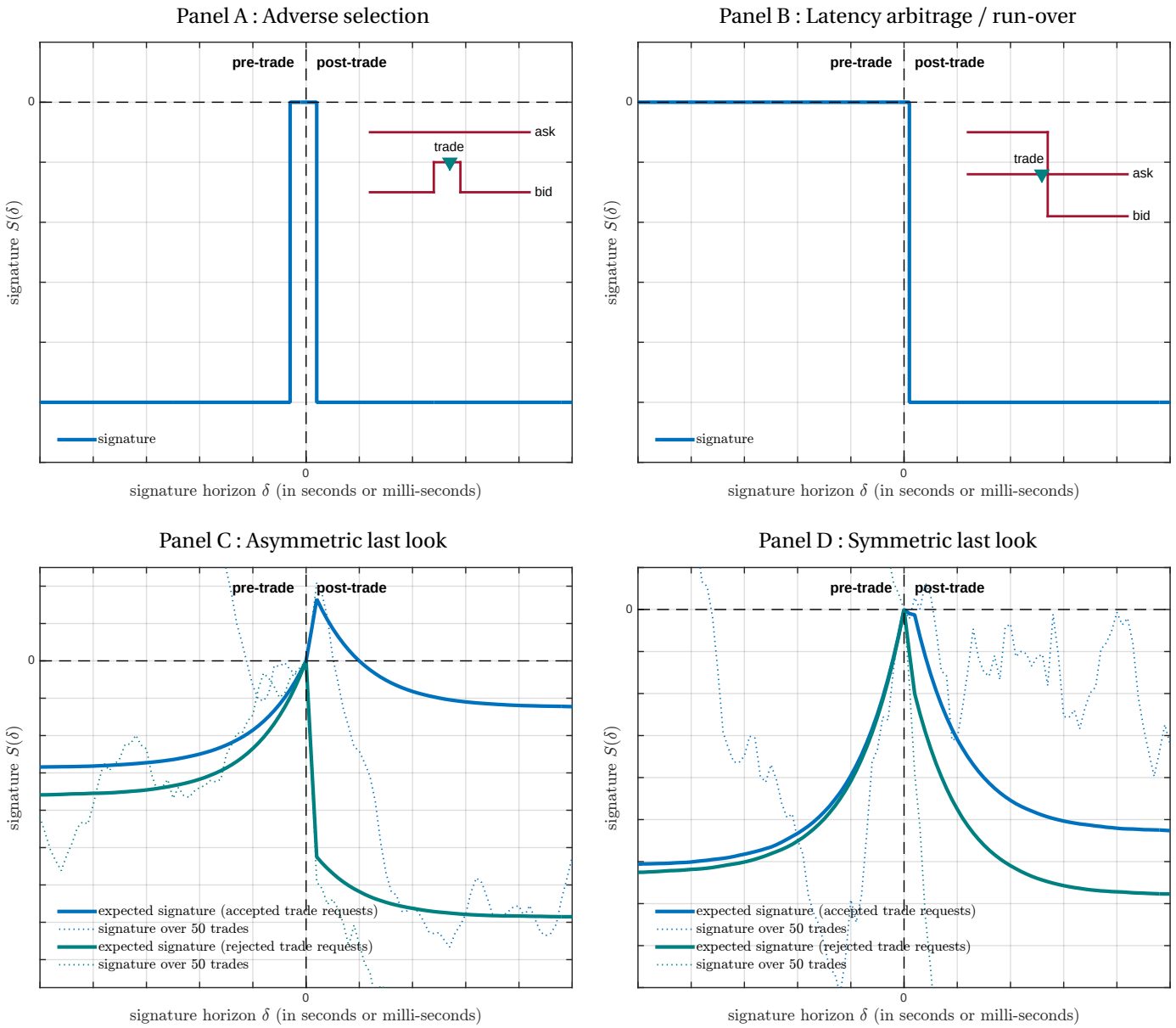
Figure 5: Signature profiles at a macroscopic level



Note. This chart provides example signature profiles associated with different trading strategies. The solid line is the expected signature for the assumed strategy and underlying price dynamics. The dotted line is a signature based on a set of 50 observations.

gressiveness of execution and associated impact towards the end of the block. It highlights the compounding nature of transaction costs for such execution strategies. Next, consider the case where the trader sets ex-ante stop-loss or take-profit levels associated with a trade. For example, they sell at a price of 1 and then hold on to the short position until either 0.8 is reached and they realise a 20% gain via a take profit buy order, or tolerate an adverse price move up to 1.10 or 10% and then exit via a stop-loss buy order. The corresponding signatures are given in Panels C and D

Figure 6: Signature profiles at a microstructure level



Note. This chart provides example signature profiles associated with different trading strategies. The solid line is the expected signature for the assumed strategy and underlying price dynamics. The dotted line is a signature based on a set of 50 observations. The signatures are drawn from the perspective of the market maker or liquidity provider.

of Figure 5. The pre-deal pattern is a consequence of the strategy – it represents the average price evolution in the run up to the exit trade conditional on reaching this point. The post-deal signature is flat, assuming that there is no alpha in the pre-defined exit level or any market impact associated with the trade.

At a microscopic level additional patterns may be encountered. Figure 6 considers the signature from a mar-

ket maker's or liquidity provider's (LP) perspective at a fine grained micro-structure resolution. In Panel A, the LP increases it's bid and tightens the spread as part of the price discovery process, but their bid quickly gets traded upon, any remaining liquidity on that side is being removed, and the price returns to its original level. It is illustrative for adverse selection incurred by the LP, and the signature has a pronounced and short-lived inverted U-shape. Panel B is closely related to this scenario, but here the LP is too slow withdrawing their bid and gets "run-over" by a latency or statistical arbitrageur that holds more accurate information regarding the future price evolution at a micro-structure level. Or it simply may be an uninformed trader that happens to require a large amount of liquidity and creates price impact by consuming all the liquidity available at that level. Finally, Panels C and D draw the signatures associated with accepted and rejected trade requests as part of an LP's "last look" trade acceptance process (see Oomen, 2017b, for further details).

2.2 Statistical properties of the signature

I now present some basic results on the signature's statistical properties under the null hypothesis that the price follows a random walk, the trader is uninformed and executes at regular periodic intervals without market impact. Studying the signature under an alternative hypothesis is difficult given the diversity of alternatives and sensitivity to specific assumptions one would need to make regarding the strategy's performance. The null hypothesis, however, provides a number of useful insights that set the stage for the remainder of this paper and also serves as a useful benchmark to test whether or not systematic patterns of any kind are observed.

Proposition 1 (stochastic trade direction) *Assume $P_t = \sigma W_t$ where W is a Standard Brownian motion, trades occur at fixed time-intervals, $t_n = n\Delta$ for $n \in \{1,2,\ldots,N\}$, for unit amounts $q = 1$, and the trade direction d_n follows a stochastic process with $E(d_n) = \mu_d$, $corr(d_n, d_m) = \rho_d^{|n-m|}$. The signature variance is then given by:*

$$\gamma(\delta) = \frac{\sigma^2}{N}\delta + \frac{\sigma^2}{N^2}\left(\mu_d^2\psi_1(M,\delta) + (1-\mu_d^2)\psi_{\rho_d}(M,\delta)\right), \tag{2}$$

where $M = \lfloor \delta/\Delta \rfloor \wedge N$, and

$$\psi_\rho(M,\delta) = \begin{cases} \frac{1}{2}NM(2\delta - \Delta(M+1)) + \frac{1}{6}M(M+1)(2M\Delta + \Delta - 3\delta) & \rho = 1 \\ \rho(1-\rho^M)\left(\frac{N\delta}{1-\rho} - \frac{\delta+N\Delta}{(1-\rho)^2} - \frac{(\rho+1)\Delta}{(\rho-1)^3}\right) + M\rho^{M+1}\left(\frac{\Delta(N-M)+\delta}{1-\rho} - \frac{2\Delta}{(1-\rho)^2}\right) & \rho \neq 1 \end{cases}. \tag{3}$$

Proof The result here is a special case of Proposition 2 below. ■

Eq. (2) shows that when the signature is calculated over a horizon that is shorter than the trader's execution frequency (i.e. $\delta < \Delta$), the second term on the right hand side of Eq. (2) drops out because $M = 0$, and the signature's volatility is characterised by the "square-root rule" in number of observations and horizon: it is proportional to

$\sqrt{\delta/N}$. This is because the price paths that enter into the signature's construction are independent Brownian motions. However, as the signature horizon expands beyond the execution frequency, the price paths start to overlap (as illustrated in Figure 3), dependence across price paths is introduced, and the convergence properties of the signature are altered. A noteworthy exception to this is when $\mu_d = \rho_d = 0$. Then the randomness in trade sign breaks the dependence accounted for by any price path overlap.

To further illustrate the signature's statistical properties, I specify $d_n \in \{1,-1\}$ to be a Markov process with state transition probability matrix:

$$P = \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}. \tag{4}$$

This process conforms to the assumptions made in Proposition 1 (see Appendix A for a proof), i.e. it has a mean of μ_d and a k-th order autocorrelation of ρ_d^k when:

$$\alpha = \frac{1}{2}(1-\rho_d)(1-\mu_d) \qquad \text{and} \qquad \beta = \frac{1}{2}(1-\rho_d)(1+\mu_d). \tag{5}$$

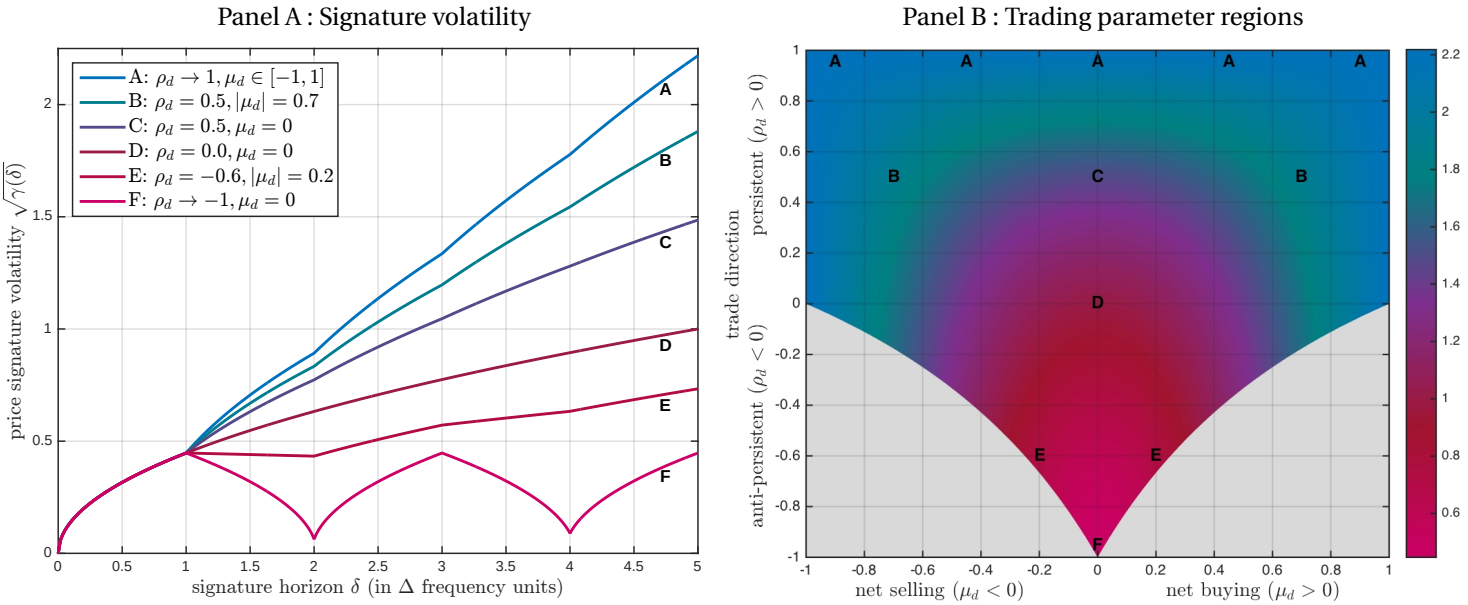
To ensure $\alpha, \beta \in [0,1]$, it is required that $\max\left(\frac{1+\mu_d}{\mu_d-1}, \frac{\mu_d-1}{1+\mu_d}\right) < \rho_d < 1$ or equivalently $\max\left(-1, \frac{1+\rho_d}{\rho_d-1}\right) < \mu_d < \min\left(1, \frac{1+\rho_d}{1-\rho_d}\right)$.

Panel A of Figure 7 draws the signature volatility as a function of horizon for different values of (μ_d, ρ_d) within the allowable parameter space, as depicted in Panel B. The special case with $\mu_d = \rho_d = 0$ is represented by scenario D where the signature's volatility grows at rate $\sqrt{\delta}$ despite the dependence in price paths for $\delta > \Delta$. The empirically more interesting scenario is where the trade sign is serially correlated. For instance, for typical algorithmic executions of large parent orders, ρ_d would be positive as the child orders normally have the same sign as the individual parent orders (even though consecutive parent orders may themselves be largely uncorrelated). In contrast, a latency arbitrage strategy on tight risk limits is expected to alternate buy and sell orders so that $\rho_d < 0$. It is this – the sign and magnitude of ρ_d – that determines to a large extent the way in which the signature volatility deviates from the baseline scenario D. When $\rho_d > 0$ the signature volatility increases due to positive dependence between overlapping price paths associated with trades of the same sign. On the other hand, when $\rho_d < 0$, successive trades in opposing direction are more likely to occur and, for those instances, any overlap in price path will cancel out and reduce the signature's volatility. In the extreme case where $\rho_d = -1$ every risk position initiated will be fully offset after Δ time-units at which point the paths start overlapping and progressively net off against one another, accounting for the repeated inverse U-shape pattern in scenario F.

Proposition 2 (stochastic trade direction and size) *Assume a setup as in Proposition 1, plus a stochastic trade size process with* $E(q) = \mu_q$, $V(q) = \sigma_q^2$, $corr(q_n, q_m) = \rho_q^{|n-m|}$. *The signature variance is then given by:*

$$\gamma(\delta) = \frac{\sigma^2}{N}\frac{\mu_q^2+\sigma_q^2}{\mu_q^2}\delta + \frac{\sigma^2}{N^2}\left(\mu_d^2\psi_1(M,\delta) + (1-\mu_d^2)\psi_{\rho_d}(M,\delta)\right) + \frac{\sigma^2}{N^2}\frac{\sigma_q^2}{\mu_q^2}\left(\mu_d^2\psi_{\rho_q}(M,\delta) + (1-\mu_d^2)\psi_{\rho_q\rho_d}(M,\delta)\right). \tag{6}$$

Figure 7: Statistical properties of the signature



Note. Panel A draws the signature volatility $\sqrt{\gamma(\delta)}$ as a function of horizon δ as given in Eq. (2), with regular trading at $\Delta = 1$ time intervals, a Markov model for the trade direction process as described in the text, and a fixed trade size $q = 1$. Panel B draws the permissible parameter region for the average trade sign μ_d and autocorrelation of trade sign ρ_d. The grey area is unachievable within the Markovian model specification. The colour coding represents the ratio of signature volatility relative to the baseline case D.

Proof See Appendix A. ■

Analysing Eqs. (2) and (6) confirms the intuition the signature's variance increases with variability in trade size.

To conclude, I illustrate how the results here address the practically relevant question of how large a sample size N is required to attain a given level of accuracy of the empirical signature statistic. Of course, this is simply answered by solving Eq. (6) for N such that the signature variance is at the required level. Consider a few examples. Assume an annualised price volatility of 20%, a sample where trades are executed every minute, a unit trade size, and a trade direction process with zero mean and autocorrelation of 75%. In this case, to attain a confidence bound around the signature of one basis point at a 5 minute horizon 2,827 observations or two days of uninterrupted trading are required (under the stated assumptions). If the signature horizon is shortened to 1 minute – the point at which dependence between price paths is eliminated – only 169 observations are required to attain the same level of accuracy. If, on the other hand, the trade size is stochastic (with $\mu_q = 1, \sigma_q = 0.4, \rho_q = 0.75$) the number of required observations increases by about 10% to 3,160. Finally, if instead of $\rho_d = 75\%$, it is negative that value, then only 225 observations are needed to attain a one basis point accuracy in the sample signature at 5 minute horizon.

This highlights that the signature properties are not only determined by the underlying price dynamics, but also intimately linked to characteristics of the trading strategy.

2.3 Functional data analysis of signatures

The theory presented in Section 2.2 provides useful intuition regarding the statistical properties of the signature, but it is based on a stylised model, imposes strong assumptions on the trading rate, and can only be used to perform analysis of a single signature under the null hypothesis of a noise trader. Furthermore, it only provides point-wise confidence bounds on the signature whereas one may be interested in testing hypotheses about the entire curve, e.g. rather than asking whether the signature is significant at a specific point, say at the 30 second horizon, often the more relevant question is whether the entire signature curve over the interval from 0 to 30 seconds is significant. This is where advances in the field of functional data analysis can be applied to analyse the signatures' statistical properties in much greater depth, under alternative hypotheses, and under much weaker assumptions.

In order to adapt the FDA methodology to the signatures, I first need to introduce some additional notation. Let $E(S(\delta)) = S^*(\delta)$, $s_n(\delta) = d_n(P_{t_n+\delta} - P_{t_n})$. Further, I assume that the set of N trades can be subdivided into K non-intersecting subsets or groups, with N_k observations in the k-th group, and $N = \sum_{k=1}^{K} N_k$. To distinguish between the groupings, all relevant variables are indexed by $k \in \{1, 2, \ldots, K\}$. Throughout the paper I also assume that $\iota' q^{(k)} = N_k$, i.e. the trade sizes determine only the relative weighting within each group, but do not affect the weighting across groups. The signature associated with the k-th grouping, can now be expressed as:

$$S_k(\delta) = \frac{1}{\iota' q^{(k)}} \sum_{n=1}^{N_k} q_n^{(k)} s_n^{(k)}(\delta). \tag{7}$$

For standard FDA to apply (such as those discussed in Zhang, 2014), I need to make two strong assumptions (aside from standard regularity conditions), namely (a) unit trade sizes or $q_n^{(k)} = 1$ and (b) the price paths that enter into the signature calculation to follow an IID Gaussian process, i.e. $s_n^{(k)}(\delta) \sim$ i.i.d. $GP(S_k^*(\delta), \gamma(\delta))$. Now a number of FDA methods are available to test – for instance – equality between group signatures:

$$H_0 : S_1^*(\delta) = S_2^*(\delta) = \ldots = S_K^*(\delta) \quad \text{for } \delta \in (0, \overline{\delta}]. \tag{8}$$

Within the context of this paper, this sort of hypothesis testing is useful to identify systematic differences in signature profiles when categorised by, e.g., currency pair, liquidity provider, trading strategy, time period, deal-size. I now list a number of FDA tests applicable when assessing the null hypothesis in Eq. (8). For an in-depth and comprehensive overview, see for instance Ramsay and Silverman (2002), Horváth and Kokoszka (2012), Zhang (2014) and references therein.

The point-wise F test (Ramsay and Silverman, 2005)

$$T_{FP}(\delta) = \frac{N-K}{K-1}\frac{SSH(\delta)}{SSE(\delta)} \sim F_{K-1,N-K} \tag{9}$$

where

$$SSH(\delta) = \sum_{k=1}^{K} N_k (S_k(\delta) - S(\delta))^2, \qquad \text{and} \qquad SSE(\delta) = \sum_{k=1}^{K}\sum_{n=1}^{N_k} (s_n^{(k)}(\delta) - S_k(\delta))^2, \tag{10}$$

denote point-wise measures of between group and within group variation.

The integrated F test (Faraway and Shen, 2004)

$$T_{FI} = \frac{N-K}{K-1}\frac{\int SSH(\delta)d\delta}{\int SSE(\delta)d\delta} \sim F_{(K-1)\kappa,(N-K)\kappa} \text{ approximately} \tag{11}$$

where $\kappa = \mathrm{tr}^2(\gamma)/\mathrm{tr}(\gamma^{\otimes 2})$, and $\gamma^{\otimes 2}(\delta_1,\delta_2) = \int \gamma(\delta_1,u)\gamma(u,\delta_2)du$. The parameter κ can be estimated from the data (see, e.g., Zhang, 2014, Chapter 5) as $\widehat{\kappa} = \widehat{\mathrm{tr}^2(\gamma)}/\mathrm{tr}(\widehat{\gamma^{\otimes 2}})$ where

$$\begin{aligned}
\widehat{\mathrm{tr}^2(\gamma)} &= \frac{(N-K)(N-K+1)}{(N-K-1)(N-K+2)}\left(\mathrm{tr}^2(\widehat{\gamma}) - \frac{2\mathrm{tr}(\widehat{\gamma}^{\otimes 2})}{N-K+1}\right), \\
\mathrm{tr}(\widehat{\gamma^{\otimes 2}}) &= \frac{(N-K)^2}{(N-K-1)(N-K+2)}\left(\mathrm{tr}(\widehat{\gamma}^{\otimes 2}) - \frac{\mathrm{tr}^2(\widehat{\gamma})}{N-K}\right).
\end{aligned}$$

An estimator of the covariance function γ is given by

$$\widehat{\gamma}(\delta_1,\delta_2) = \frac{1}{N-K}\sum_{k=1}^{K}\sum_{n=1}^{N_k}(s_n^{(k)}(\delta_1) - S_k(\delta_1))(s_n^{(k)}(\delta_2) - S_k(\delta_2)). \tag{12}$$

The L2-norm test (Zhang and Chen, 2007)

$$T_{L2} = \int SSH(\delta)d\delta \sim \beta\chi^2_{(K-1)\kappa} \text{ approximately} \tag{13}$$

where $\beta = \mathrm{tr}(\gamma^{\otimes 2})/\mathrm{tr}(\gamma)$, and can be estimated as $\widehat{\beta} = \mathrm{tr}(\widehat{\gamma^{\otimes 2}})/\mathrm{tr}(\widehat{\gamma})$.

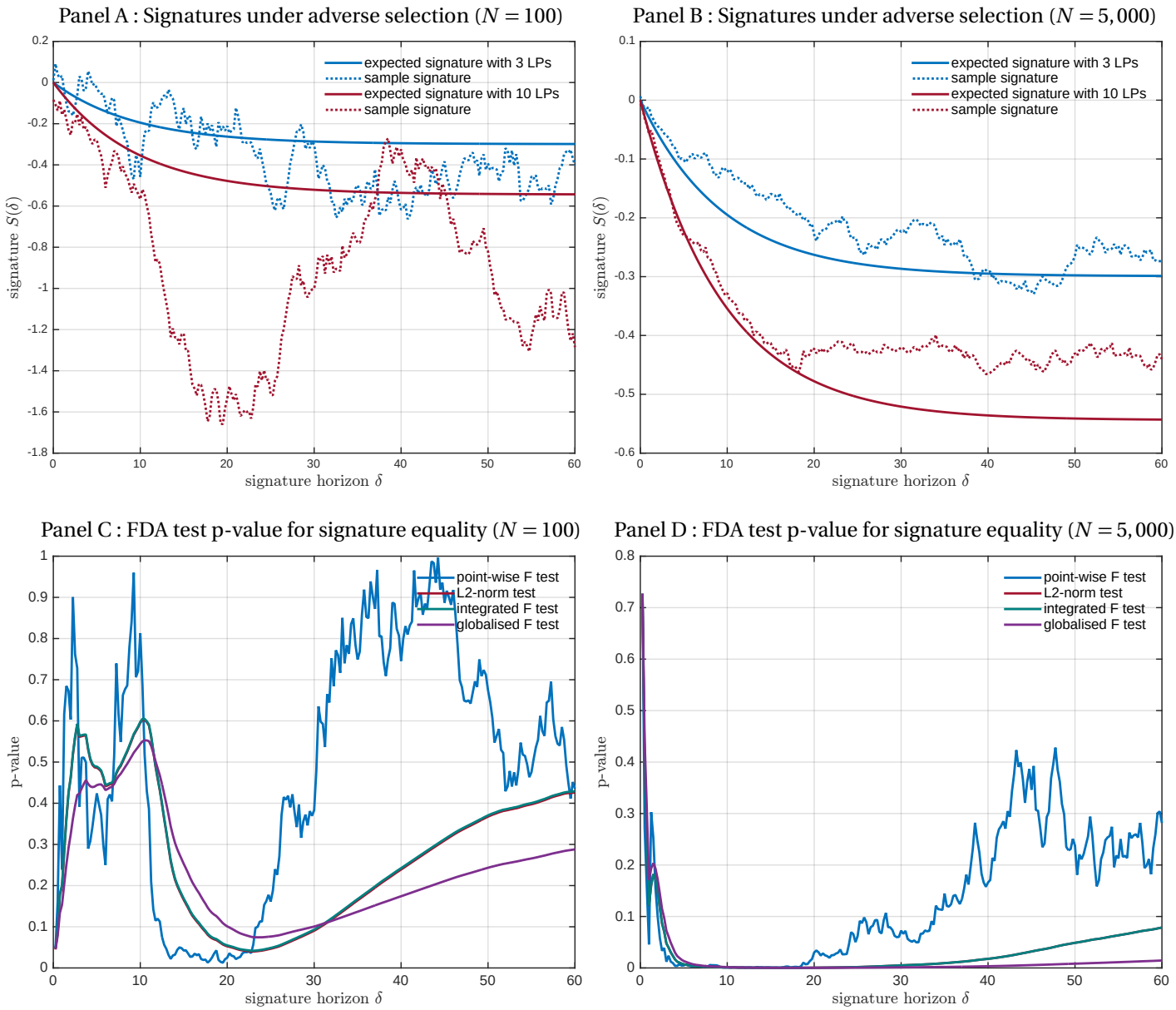
The globalised F test (Zhang and Liang, 2014)

$$T_{FG} = \int T_{FP}(\delta)d\delta \sim a\chi^2_b \text{ approximately} \tag{14}$$

where

$$a = \overline{\delta}\frac{N-K-2}{(K-1)(N-K)}\mathrm{tr}(\gamma_c^{\otimes 2}), \qquad b = \overline{\delta}^2\frac{(K-1)(N-K)^2}{(N-K-2)^2}\mathrm{tr}(\gamma_c^{\otimes 2}), \qquad \text{and} \qquad \gamma_c(\delta_1,\delta_2) = \frac{\gamma(\delta_1,\delta_2)}{\sqrt{\gamma(\delta_1,\delta_1)\gamma(\delta_2,\delta_2)}}.$$

Figure 8: FDA application to simulated signature data



Note. Panels A and B draw the expected group signatures $S_k^*(\delta)$ and sample group signatures $S_k(\delta)$ for the simulation explained in the text and different sample sizes. Panels C and D draw the corresponding FDA tests as a function of signature horizon, i.e. the point-wise F test in Eq. (9), the integrated F test in Eq. (11), the L2-norm test in Eq. (13), and the globalised F test in Eq. (14).

An extensive comparison of these and numerous other FDA tests is provided by Górecki and Smaga (2015) using real and simulated data. They find that the above four tests are amongst the best performing across a variety of scenarios.

To demonstrate how these tests can be applied to signature analysis, I set up a simple simulation that conforms

with the assumptions made thus far (i.e. unit trade size, and an IID Gaussian process). I produce two group signatures with a mean $S_k^*(\delta)$ for $k \in \{1,2\}$ and $\delta > 0$ that matches the adverse selection profile experienced by an LP that is providing liquidity into an aggregator within the model of Oomen (2017a).[7] The first group corresponds to trades obtained from an aggregator where the LP is in competition with two other LPs. The second group constitutes trades obtained from an aggregator where a total of ten LPs compete for the trader's flow and the adverse selection and signature decay is therefore stronger. Panels A and B of Figure 8 draw the expected group signatures $S_k^*(\delta)$ together with the sample group signatures $S_k(\delta)$ calculated from a set of $N_k = 100$ and 5,000 trades respectively. As expected, the signatures based on the smaller set of trades are less accurate than those based on the larger set of trades. Panels C and D draw the p-values of the various FDA tests for the hypothesis that $S_1^*(\delta) = S_2^*(\delta)$ for $\delta \in [0, \overline{\delta}]$. For the smaller sample, both the point-wise F test and the composite tests for the entire signature curve cannot firmly reject the null hypothesis at any signature horizon. There is marginal point-wise significance between horizons of about 10 and 20 where S_2 deviates more from the expected signature than elsewhere on the interval. But the composite tests over the full signature curve have p-values of around 30%-40% and hence cannot identify a statistically significant difference between the two signatures. For the larger sample size, the p-values are substantially smaller as expected. The global F-test firmly rejects the hypothesis of equal signatures across all horizons except for the very shortest ones. This is in contrast to the point-wise F test which cannot distinguish between the signatures beyond a horizon of $\delta = 20$ or so.

2.4 Functional data analysis of signatures using bootstrap methods

The FDA tests above require two strong assumptions which in the context of signature analysis are unlikely to be satisfied, i.e. unit trade sizes and IID-ness. In practice, trade sizes vary as a function of numerous factors (e.g. liquidity demand and availability, strength of signal or conviction, cost of trading) and a signature normalised by volume is then a natural quantity of interest. Additionally, IID-ness would routinely be violated due to potential overlap in signature price paths and the fact that trade signs tends to be correlated over time (e.g. due to order splitting and staggered executions, varying reaction times to macro economic news events, herding effects). Specifically, I am concerned here with dependence in the signed price paths that feed into the construction of the signature. Even when the price process is a martingale, such dependence arises somewhat "mechanically" as soon as the signature horizon exceeds the time-displacement of consecutive trades, i.e. $s_n(\delta)$ and $s_m(\delta)$ share an identical portion of the

[7] The post-trade adverse selection within this model is given as $-(1-\beta^\delta)\omega\sqrt{1-\rho}\,c_M$ where δ is the signature horizon, β measures the dependence in the LP's measurement error of the unobserved efficient price, ρ denotes the correlation of measurement error across LPs, ω is the efficient price volatility, and c_M is a constant that depends on the number of LPs M that compete in the aggregator. In the simulations I set $\sigma = 0.5, \beta = 0.9, \omega = 0.5, \rho = 0.5$ and $M = 3$ or 10.

price path when $|t_n - t_m| < \delta$ and this gives rise to correlation between $s_n(\delta)$ and $s_m(\delta)$ unless it is broken by independence of the trade signs d_n and d_m.[8] The reason why the standard tests are invalidated in those scenarios can be understood by considering the SSE in Eq. 10 (set $K = 1$ for simplicity, a similar intuition applies to $\widehat{\gamma}$ in Eq. 12). When the data is IID, the SSE is an unbiased estimator of $V(S(\delta))$. But when dependence in the signature's constituent signed price paths $s_n(\delta)$ is introduced, the SSE can become a severely biased estimator of the variance because it ignores all the covariance terms. In such situations, one would normally consider a HAC-type bias correction but what is unusual here is that the dependence structure of s_n is known conditional on (t_n, d_n). To make this explicit, assume the price P follows a Brownian motion with variance σ^2, set $K = 1$ and $q'\iota = N$ merely for simplicity of notation, and condition on the trade time t_n, trade sign d_n, and trade size q_n. Then:

$$V(S(\delta)|t,d,q) = \lambda_\omega(\delta)\frac{\delta\sigma^2}{N}, \tag{15}$$

where $\omega_n = q_n d_n$ and

$$\lambda_\omega(\delta) = \frac{1}{\omega'\omega\delta}\begin{pmatrix}\omega_1\\ \omega_2\\ \vdots\\ \omega_N\end{pmatrix}'\begin{pmatrix}\delta & |t_1+\delta-t_2|^+ & \cdots & |t_1+\delta-t_n|^+\\ |t_1+\delta-t_2|^+ & \delta & \ddots & \vdots\\ \vdots & \ddots & \ddots & |t_{N-1}+\delta-t_N|^+\\ |t_1+\delta-t_n|^+ & \cdots & |t_{N-1}+\delta-t_N|^+ & \delta\end{pmatrix}\begin{pmatrix}\omega_1\\ \omega_2\\ \vdots\\ \omega_N\end{pmatrix}. \tag{16}$$

Note that when there is no overlap in the signature data (i.e. $\min_n(t_n - t_{n-1}) > \delta$) then the above matrix is diagonal and $\lambda_\omega(\delta) = 1$. With overlap, however, the function $\lambda_\omega(\delta)$ provides a multiplicative bias correction for the IID variance estimator (as in Eq. 16) and more generally constitutes a useful measure to gauge the impact of dependence in the signature data on the validity of the FDA tests that rely on IID-ness. For instance, $\lambda_\iota(\delta)$, i.e. where $\omega_n = 1$, measures the aggregate amount of time overlap in the signature price paths. Consequently, the difference between $\lambda_{dq}(\delta)$ and $\lambda_\iota(\delta)$ is a measure that isolates the dependence introduced or destroyed by the variability of trade signs and sizes. Finally, $N/\lambda_{dq}(\delta)$ informally represents the number of "effective" observations in the data set.

When the assumptions underlying the FDA tests are violated, the statistic itself may still be a useful metric but its p-values cannot be reliably obtained anymore from the distribution theory. For instance, when there is within group data dependence we know that the SSE can be distorted but the SSH – and therefore the L2-norm FDA test – may still be a valid metric. Similarly, if the dependence is such that it merely introduces a bias in the SSE that is

[8] The signatures themselves, on the other hand, are independent from one another provided that they are calculated over non-intersecting time periods displaced by more than δ and the price process is a martingale. Functional data analysis when there is dependence in the underlying price process (rather than in the set of sampled price paths) is studied by, for instance, Horváth, Kokoszka, and Rice (2014); Zhang (2016).

uniform across groups then again the FDA tests listed above are useful quantities for statistical inference, except that confidence bounds and critical values need to be adjusted.[9] This can be done using bootstrap resampling methods. The general approach is as follows. Let $\{X_n\}_{n=1}^N$ denote a set of observations. From this, B bootstrapped samples can be obtained as follows:

- randomly draw a series of indices $(i_{b,1}, i_{b,2}, \ldots)$ from a uniform distribution over $\{1,2,\ldots,N\}$,
- determine a series of associated block lengths $(l_{b,1}, l_{b,2}, \ldots)$,
- the b-th bootstrap sample is then obtained as the first N elements of the series,

$$\widetilde{X}_b = (X_{\iota(i_{b,1}:i_{b,1}+l_{b,1}-1)}, X_{\iota(i_{b,2}:i_{b,2}+l_{b,2}-1)}, \ldots),$$

  where $\iota(i) = i - \lfloor \frac{i-1}{N} \rfloor N$, $n:m = (n, n+1, \ldots, m)$, and $X_{n:m} = (X_n, X_{n+1}, \ldots, X_m)$,
- repeat these steps for all $b \in \{1,2,\ldots,B\}$.

The above procedure describes the IID bootstrap (Efron, 1979) when the block length $l = 1$, the moving block bootstrap (Künsch, 1989) when l is set to a fixed positive integer c and sampling of i_b is done over $\{1,2,\ldots,N-c+1\}$, the circular block bootstrap (Politis and Romano, 1992) when l is a fixed positive integer, and the stationary bootstrap (Politis and Romano, 1994) when l is randomly drawn from a Geometric distribution with positive mean.

In the context of this paper, $X_n = (s_n^{(k)}(\delta), d_n^{(k)}, q_n^{(k)})$ for $n \in \{1,2,\ldots,N_k\}$, $k \in \{1,2,\ldots,K\}$ and bootstrapped samples of the signature data, denoted by $\{\tilde{s}_{b,n}^{(k)}(\delta), \tilde{d}_{b,n}^{(k)}, \tilde{q}_{b,n}^{(k)}\}_{b=1}^B$, can be obtained via the above procedure. From this, compute:

$$\widetilde{SSH}_b(\delta) = \sum_{k=1}^K \widetilde{Q}_b^{(k)} \left(\widetilde{S}_{b,k}(\delta) - \widetilde{S}_b(\delta)\right)^2, \tag{17}$$

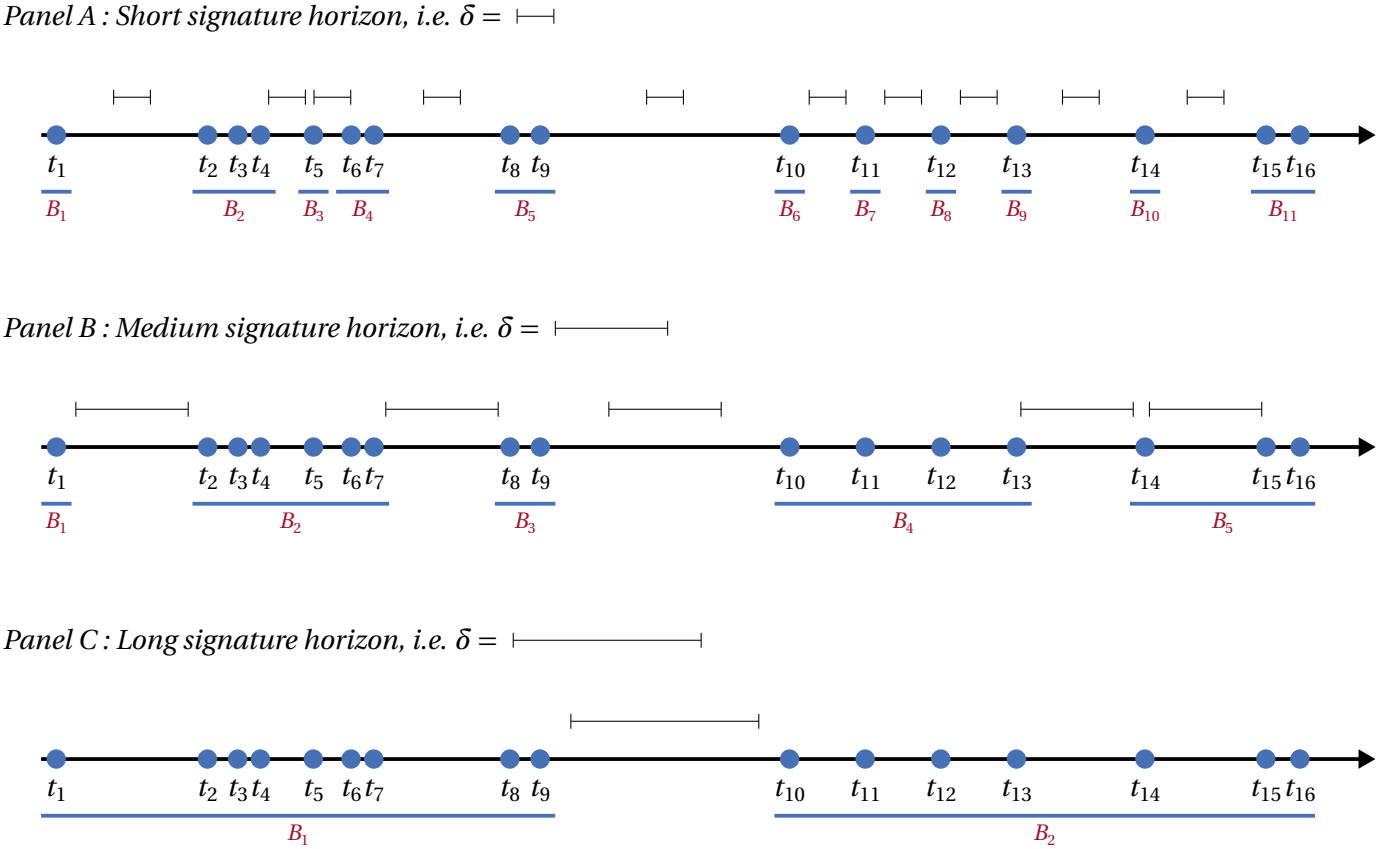
$$\widetilde{SSE}_b(\delta) = \sum_{k=1}^K \sum_{n=1}^{N_k} \left(\tilde{q}_{b,n}^{(k)} \tilde{e}_{b,n}^{(k)}(\delta)\right)^2, \tag{18}$$

where

$$\widetilde{S}_{b,k}(\delta) = S_k(\delta) + \frac{\sum_{n=1}^{N_k} \tilde{q}_{b,n}^{(k)} \tilde{e}_{b,n}^{(k)}(\delta)}{\widetilde{Q}_b^{(k)}} \qquad \text{and} \qquad \widetilde{S}_b(\delta) = \frac{\sum_{k=1}^K \widetilde{Q}_b^{(k)} \widetilde{S}_{b,k}(\delta)}{\sum_{k=1}^K \widetilde{Q}_b^{(k)}},$$

[9] With non-IID data, the decision whether to apply the L2-norm test or say the integrated F test is similar to whether one would bootstrap the sample mean or the t-statistic. Typically the former is the more natural one, but depending on the data properties both may be valid. Also note that the weighting by N_k of SSH in Eq. (10) also relies on the IID assumption, and may require adjustments particularly when there is significant between group dependence, or strong heterogeneity in trade arrival frequency across groups. This is beyond the scope of this paper, and of no concern for the empirical analysis below.

Figure 9: Illustration of the adaptive block bootstrap construction



Note. This chart illustrates the construction of the adaptive blocks as a function of signature horizon δ for use in the bootstrap resampling procedure. A new block is initiated at t_n iff $t_n - t_{n-1} > \delta$. The identified blocks are denoted by B_i and underlined as ——.

and $\widetilde{Q}_b^{(k)} = \sum_{n=1}^{N_k} \tilde{q}_{b,n}^{(k)}$, $\tilde{e}_{b,n}^{(k)}(\delta) = \tilde{s}_{b,n}^{(k)}(\delta) - S_k(\delta)$. Next, bootstrap samples of the FDA tests $(\widetilde{T}_{b,FP}, \widetilde{T}_{b,FI}, \widetilde{T}_{b,L2}, \widetilde{T}_{b,FG})$ can be computed by substituting the bootstrapped SSH and SSE values from Eqs. (17) and (18) into Eqs. (9), (11), (13), and (14). Bootstrapped p-values of the sample FDA test statistic can be computed as:

$$\frac{1}{B}\sum_{b=1}^{B} 1_{T_* < \widetilde{T}_{b,*}}. \tag{19}$$

Similarly, confidence bounds on the signature itself can be obtained from the set of bootstrap signature samples $\{\widetilde{S}_{b,k}(\delta)\}_{b=1}^{B}$ as the $p-$th percentile, denoted by $\widetilde{S}_k^{(p\%)}(\delta)$. Sometimes it is useful to condense aspects of the signature distribution into a single number by calculating integrated signature percentiles, i.e.

$$\mathcal{I}_{\underline{\delta},\overline{\delta}}^{(p\%)} = \frac{1}{\overline{\delta} - \underline{\delta}} \int_{\underline{\delta}}^{\overline{\delta}} \widetilde{S}_k^{(p\%)}(\delta) d\delta. \tag{20}$$

The intuition obtained from the simple model in section 2.2 is helpful to inform the precise bootstrap implementation. For instance, the IID bootstrap is only expected to provide reliable results when the signature samples $s_n(\delta)$ are uncorrelated. A sufficient condition is for the trade signs d_n to be serially uncorrelated or for the signature horizon to be shorter than the minimum trade duration in the sample, i.e. $\delta < \min_n(t_n - t_{n-1})$. As discussed, in practice neither are likely to be satisfied and so the bootstrap design will need to account for the dependence in the data. The stationary bootstrap described above is a good candidate because it is designed for that purpose. The mean of the Geometric distribution from which the block sizes are drawn can be selected via the procedure of Politis and White (2004) and Patton, Politis, and White (2009) using as input the strength of correlation in the time series. Note, however, that this procedure is agnostic to the actual event times t_n which are known to be a key driver of dependence in the signature data. This suggests an alternative bootstrap design. In particular, for given a signature horizon δ and the set of time stamps t_n, it is straightforward to identify blocks of observations that are mutually uncorrelated simply by virtue of absence of overlap in time. Therefore, instead of selecting blocks that start at random points in the sample as in the stationary bootstrap, blocks can be determined on the basis of t_n and δ, i.e. define the set of indices $\{i_{\delta,j}\}_{j=1}^{n_\delta}$ that satisfies:

$$t_{i_{\delta,j}+1} - t_{i_{\delta,j}} \geq \delta \qquad \text{for } j = 1,2,\ldots,n_\delta \qquad \text{and} \qquad t_{n+1} - t_n < \delta \quad \text{for all } n \notin \{i_{\delta,j}\}_{j=1}^{n_\delta}, \tag{21}$$

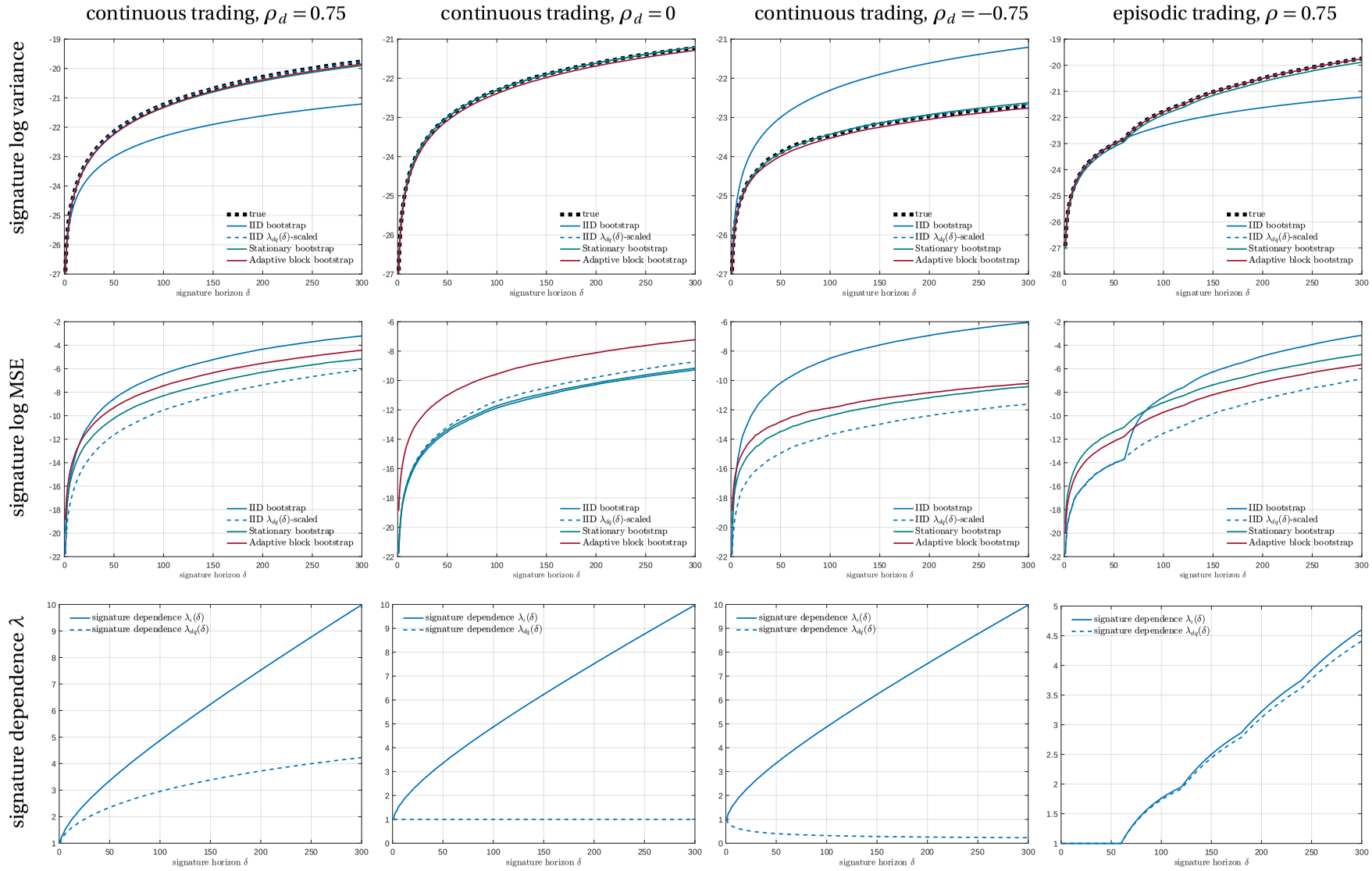
and define the associated $n_\delta + 1$ blocks as:

$$\overline{X}_j = X_{i_{\delta,j}+1}, \ldots, X_{i_{\delta,j+1}} \qquad \text{for } j = 0,2,\ldots,n_\delta \tag{22}$$

where $i_{\delta,0} = 0$ and $i_{\delta,n_\delta+1} = N$. Because the blocks perfectly preserve the dependence structure of the data, and the blocks are mutually uncorrelated, the IID bootstrap can now be applied to $\overline{X}$. I will refer to this method as the "adaptive block bootstrap" to reflect that the selection of blocks adapts to the signature horizon. Figure 9 provides an illustration of the adaptive block construction.

I conclude this section with a small simulation study to gauge the performance of the alternative bootstrap methodologies, focussing on the (a) IID bootstrap, (b) the stationary block bootstrap, and (c) the adaptive block bootstrap. On the basis of simulated signature data $\{t_n, s_n(\delta), d_n, q_n\}_{n=1}^N$, I judge the performance of the bootstrap methods by their ability to accurately estimate the signature variance. The simulation design is informed by the empirical data analysed in the next section to ensure it reflects some salient properties of signature data encountered in practice. I set $N = 1,000$, $B = 10,000$, and without loss of generality $K = 1$. The logarithmic price path P_{t_n} is simulated by a Brownian motion. The trade sign d_n follows the Markovian model described in Eqs. (4 – 5) with $\mu_d = 0$ and I then distinguish between scenarios with positive, zero, and negative serial correlation, i.e. $\rho_d \in \{-0.75, 0, 0.75\}$.

Figure 10: Simulation study



Note. This chart reports simulation results to assess the performance of the IID, the $\lambda_{dq}(\delta)$-scaled IID, stationary, and adaptive block bootstrap procedures for estimating the signature variance under alternative assumptions on the trading process (in columns, with design as described in the text). The top row draws the true signature variance as a function of δ together with the bootstrap estimates averaged across 10,000 simulation replications. The middle row draws the corresponding mean-squared error of the bootstrap estimates. The bottom row draws the function $\lambda_\iota(\delta)$ (solid line) and $\lambda_{dq}(\delta)$ (dashed line) given by Eq. (16) as measures of dependence in the signature data.

The trade size q_n is stochastic and simulated using the below model

$$q_n = \omega + \beta q_{n-1} + \alpha q_{n-1}\varepsilon_n^2, \tag{23}$$

where $\varepsilon_n \sim$ i.i.d. $\mathcal{N}(0,1)$. For this process to attain properties characterised by $(\mu_q, \sigma_q^2, \rho_q)$ as in Proposition 2, the parameters in Eq. (23) need to be set via the following mapping (Posedel, 2005):

$$\alpha = \frac{\sigma_q\sqrt{2\mu_q^2+3\sigma_q^2-2\rho_q^2(\sigma_q^2+\mu_q^2)}-\rho_q\mu_q^2}{2\mu_q^2+3\sigma_q^2}, \quad \beta = \rho_q - \sigma_q\sqrt{(1-\rho^2)/(2\sigma_q^2+2\mu_q^2)}, \quad \text{and} \quad \omega = (1-\alpha-\beta)\mu_q.$$

I set $\mu_q = 1$, $\sigma_q = 0.35$, and $\rho_q = 0.65$ which generates a substantially asymmetric distribution of trade sizes as commonly observed in practice. Finally, the trade time stamps t_n are simulated by two different methods. The first simulates the durations between two consecutive time stamps as squared geometric IID random variables with mean $\Delta = 5$ time units. This corresponds to a continuous trading session with irregularly spaced trades and the occasional long period of inactivity.[10] The second method, simulates the time stamps as $t_n = \Delta \mod (n-1,20) + \lfloor \frac{n-1}{25} \rfloor \overline{\delta}$ where $\Delta = 5$. With $N = 1,000$ this constitutes a series of 50 independent trading sessions where, within each, 20 trades are executed consecutively at fixed intervals. This reflects an episodic trading scenario where – for instance – a trader executes one large sized trade every morning and afternoon for one month and uses an algorithm for this that splits each parent order in 20 tranches.

Figure 10 reports the results over 10,000 independent simulation replications, for signature horizons between $\underline{\delta} = 0, \overline{\delta} = 300$. To gauge any biases, the top row reports the true and bootstrapped logarithm of the signature variance. To assess efficiency of the estimators, the middle row reports the mean squared error (MSE). And as a measure for the degree of dependence in the signature data, the functions $\lambda_\iota(\delta)$ and $\lambda_{dq}(\delta)$ are drawn in the bottom row. Starting with the continuous trading scenarios (in the first three columns), I find that the stationary and adaptive block bootstrap methods produce essentially unbiased results. As anticipated, the IID bootstrap is severely biased when there is serial dependence in the trade sign. Interestingly, when multiplying the IID bootstrap signature variance estimates by $\lambda_{dq}(\delta)$, the bias is removed. Also, judging from a MSE perspective, the scaled IID bootstrap is the most efficient one when $\rho_d \neq 0$ (when $\rho_d = 0$, the correction is not needed and so it will merely add measurement error to the estimator, hence the higher MSE). It should be noted, however, that while for the specific purpose of signature variance estimation the scaling method works very well in this simulation design, it is not easily extensible to more general statistics like the FDA tests and hence I do not elaborate on it. Turning to the adaptive bootstrap, the results indicate it is unbiased but tends to have a higher MSE than the alternatives. This is explained by the block

[10] Assuming a 5.5 day / 24 hour trading session, this translates into an average duration between trades of about 8 minutes, but regularly going as low as 10 seconds and as high as 2 hours.

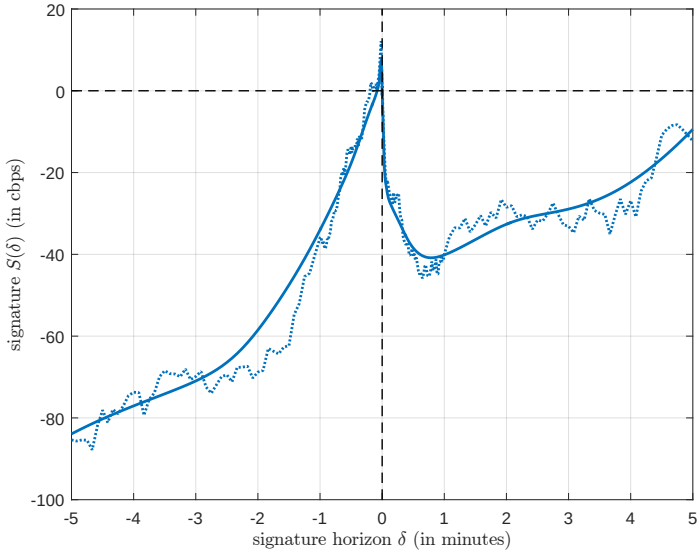
size selection. The adaptive block length selected is on average 44.4 observations with a standard deviation of 43.7 whereas the stationary bootstrap selects substantially shorter blocks with an average (standard deviation) of 24.1 (6.2) when $\rho_d = 0.75$, 1.6 (0.9) when $\rho_d = 0$, and 29.1 (13.3) when $\rho_d = -0.75$. The adaptive block bootstrap is agnostic to the trade sign correlations and is in that sense quite conservative. With larger blocks selected, there is less randomness in the bootstrap samples and the efficiency of the estimator reduced. However, with episodic trading, the results are quite different. While the both the stationary and adaptive block bootstrap yield unbiased estimates, the latter now achieves the lowest MSE (scaled IID aside). Every block it selects corresponds to an episode of exactly 20 consecutive trades whereas the stationary bootstrap selects an average block length of 63.0 with a standard deviation of 3.7. Considering that by design the dependence in the signature data does not extent beyond 20 observations, the longer than required block length for the stationary bootstrap method translates into an efficiency loss.

In summary, both the stationary and adaptive block bootstrap methods produce good first order results. They capture the dependence in the data and the resulting estimates of the signature variance are largely unbiased. Depending on the specific simulation design, one method may outperform the other in term of efficiency, with a more conservative block length selection generally translating in a reduction of efficiency as measured by the MSE.

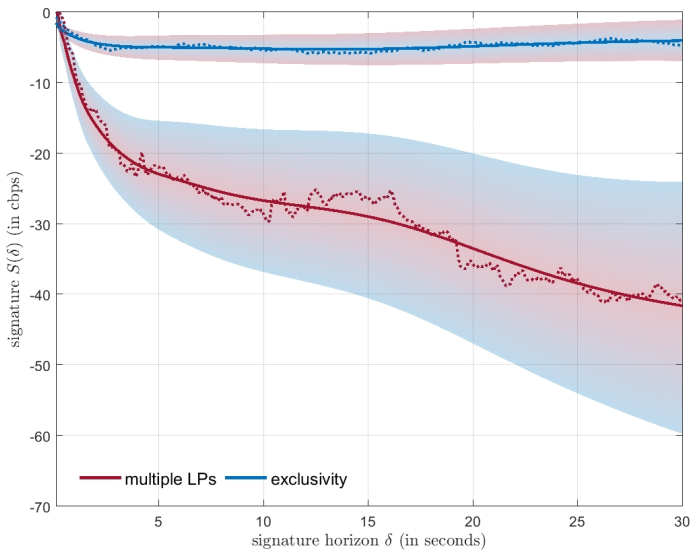
These findings suggest a hybrid bootstrap methodology that combines the block selection procedure of the stationary and adaptive block bootstraps. Both are found to select blocks conservatively because they either only consider the time overlap of price paths but not the serial dependence in the trade sign or vice versa. Specifically, the stationary bootstrap can over-estimate the block lengths required if the price paths are displaced by a sufficient amount to make them independent regardless of trade sign correlation while the adaptive bootstrap can over-estimate the block lengths when the trade sign correlation decays faster than the time-overlap in price paths. It is natural to combine the two and use all information available regarding the dependence structure of the signature data by – for instance – taking the adaptive block selection procedure as a baseline but then to impose a maximum block length as determined by the stationary bootstrap. For the case with $\rho_d = 0$ in the simulations above, such a procedure would approximate the efficiency of the IID bootstrap, when $\rho_d \neq 0$ it would prevent the overly conservative adaptive block selection, and for the episodic trading scenario it would yield equivalent performance to the best performing method. It would also address the edge case where $t_n = n\Delta$ (i.e. a single uninterrupted regular sequence of observations) and the adaptive block procedure fails when $\delta > \Delta$ by identifying the entire sample as one block, invalidating the subsampling approach. Because none of these scenarios are encountered in the case studies below, and the stationary and adaptive block bootstrap procedures give qualitatively identical results, I will not elaborate on the hybrid procedure but expect it to be an important refinement when applied across a variety of

Figure 11: Aggregation versus LP exclusivity

Panel A : Macro signature for baseline scenario



Panel B : Post-deal micro signature by aggregator setup



Note. Panel A draws the macro signature for the base line scenario where the trader executes in an aggregator with multiple LPs. Panel B draws the post-deal micro signature up to a horizon of 30 seconds for trades executed via the aggregator with "multiple LPs" and separately for the case where the trader deals with a single LP under an "exclusivity" arrangement. 95% confidence bounds are obtained via the adaptive bootstrap procedure using $B = 10,000$ replications. Source data provided by Deutsche Bank.

heterogenous datasets.

3 Case studies

I now apply the FDA and resampling methods discussed in Sections 2.3 and 2.4 to a number of practical case studies using data from a live trading environment in the currency spot market. All client and counter-party names are anonymised, and the time period and currency pair(s) are not reported. A common thread throughout, is that each study considers a trader who executes electronically using an aggregator that consolidates liquidity provided by competing LPs, but for different reasons contemplates a change to their execution setup. Within this context, I show that the signature analysis allows the trader to test the theoretical prediction made by work on optimal aggregator design (as discussed in Oomen, 2017a; Butz and Oomen, 2018), and as such provides a useful bridge between theory and practice that can be used for data driven decision making and performance evaluation. For reference throughout, Table 4 and Figure 16 provide various summary statistics of the case study data and associated bootstrap parameters.

3.1 Aggregation versus LP exclusivity

In this case study, I consider a trader that manages its risk position by hedging into an aggregator. The design of the aggregator has evolved over time, and in particular the number of participating LPs has gradually increased. But this has come with more frequent rejects of trade requests, more conservative spreads quoted by individual LPs, and a resource intensive relationship management process. Hence, the trader is open to alternative setups with the aim of improving liquidity access and operational efficiency.

Panel A of Table 4 shows that the trader executes frequently but in relatively small size per individual ticket. Nonetheless, their trading activity is characterised by a very pronounced pre- and post-deal signature pattern as illustrated by Figure 11. In particular, the post-deal micro signature in Panel B (red line) is indicative of latency sensitive and directional flow: it exhibits a sharp decay shortly after trade inception, and continues to deteriorate up to a horizon of one minute with a subsequent partial reversion. The superimposed (adaptive block) bootstrapped confidence bounds indicate that the signature pattern is highly significant with an integrated impact measure $\mathcal{I}_{30s}^{50\%}$ of about −30 centi-basis points (cbps) and a much more conservative measure $\mathcal{I}_{30s}^{95\%}$ that is still substantially negative at −18 cbps. These systematic price dynamics reduce the LPs' ability to internalise flow: the market is likely to move against the LP before they are able to work out of the trade (e.g. Butz and Oomen, 2018, show that typical internalisation horizons are measured in multiple minutes for liquid currency pairs and tens of minutes for less liquid ones).

An important question is whether the signature pattern reflects the intrinsic nature of the trader's flow or whether instead it is broadly benign at source but that adverse effects are introduced as part of the aggregation process. The answer matters because if the trader's flow is intrinsically latency sensitive and directional, then there is little that can be done about the last look rejects and LPs' defensive pricing. If, however, the aggregator design is the source of these effects, then straightforward changes can improve matters. To measure the effects of aggregator design on the observed price signature, the trader conducts an experiment and temporarily moves execution of a high volume currency pair to a single LP who manages the flow by passively internalising the risk at 100% fill rate and offers a competitive spread in view of the exclusivity arrangement.

Panel B of Figure 11 contrasts the post-deal micro signature of the trader's flow for the baseline scenario where they execute against multiple LPs (red line) with that where they execute against a single LP (blue line). The confidence bounds indicate a highly significant difference in signature between the two execution strategies. Also, all FDA tests for equality between the two signatures are firmly rejected (unreported). These results demonstrate that the trader's flow is intrinsically benign and that the previously observed sharp signature decay is largely due to the aggregator design and the risk management approach adopted by at least some of the LPs. Reducing the number of

LPs in the aggregator – to a single one in this case, carefully selected based on their ability to internalise – results in a meaningful reduction of execution costs for the trader, a simplification of workflows, and reduced overheads on LP relationship management.

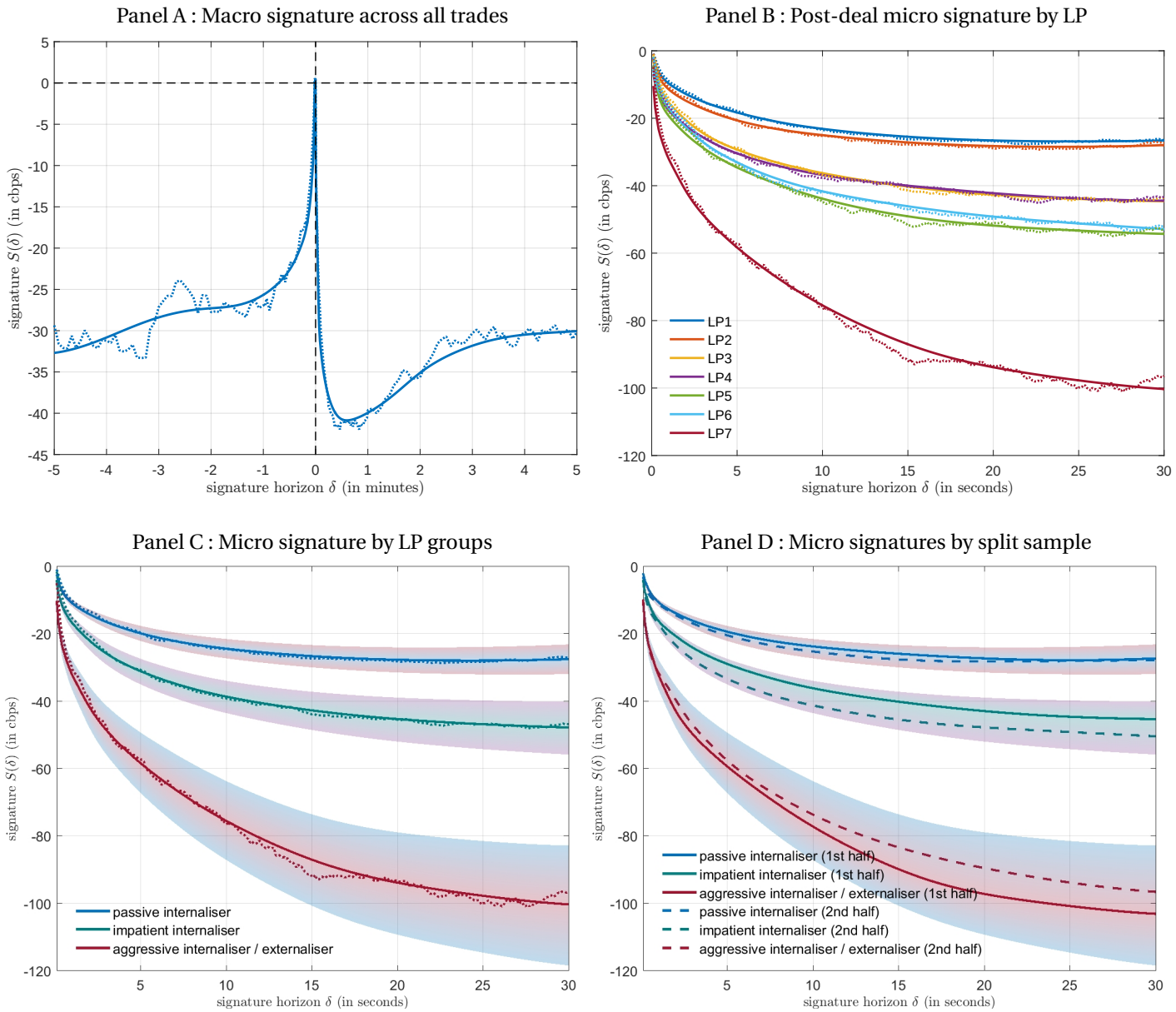
A few concluding remarks from a purely statistical perspective. An indication of the impact that data dependence has on the signature properties, and the importance of correctly accounting for it, can be gleaned from Table 4. For instance, at a thirty second horizon the signature's confidence bounds in the LP exclusivity scenario are more than double the naively computed ones that ignore any dependence due to overlap in price paths and trade sign correlation, i.e. the λ^{30s}_{tdq} measure is 2.1. Also worth highlighting is the very different behaviour of the stationary and adaptive block bootstrap procedures. The stationary bootstrap selects an average block length of 298 observations due to the very strong and persistent autocorrelation of trade signs (see top right panel in Figure 16). The adaptive block bootstrap, on the other hand, considers only the trade time stamps and ensures a sufficient displacement between consecutive price paths to break dependence which for this data set is already achieved at an average block length of only 5 observations. But the largest individual block selected by this procedure is 888 observations spanning nearly 30 minutes corresponding to an episode where a burst of trades were executed in quick succession. Despite these differences, both bootstrap methods produce qualitatively equivalent results though the efficiency of the adaptive blocks is expected to be superior for this data.

3.2 Monitoring of LP risk management style

In this case study, I consider a trader who sources liquidity from seven LPs across a range of currency pairs. The distribution of trade sizes won by each LP are similar (e.g. a median of 1,000,000 base notional) although some LPs win a substantially larger proportion of the total flow than others, see Panel B of Table 4. The trader wants to better understand the liquidity offered by the LPs, motivated by some anecdotal evidence that indicates some executions have substantially higher market impact than others. To investigate, I analyse a recent history of the full set of the trader's executions together with associated anonymised LP identifiers.

Panel A of Figure 12 draws the macro signature across the trader's aggregate activity with all LPs. It shows a pronounced "take profit" like pre-deal pattern, an overall "reversal" pattern, and a very strong post-deal signature decay over a short horizon. Panel B concentrates on the post-deal micro signature, grouped by LP identifier. It reveals a large dispersion in post-deal impact across LPs. To assess whether these differences are significant, I perform a series of pairwise FDA tests for equality between signatures. Table 1 reports the p-values for the L2-norm test (the results for the other FDA tests are numerically very similar and qualitatively identical, see Table 5 in Appendix B). It suggests a natural grouping of LPs into those with very modest decay of about 25 cbps (LPs 1 & 2), those with mod-

Figure 12: Monitoring of LP risk management style



Note. This chart draws the trader's macro signature pre- and post-deal over a five minute horizon in Panel A, and the post-deal micro signature over the first 30 seconds and grouped by LP identifier in Panel B. Panel C draws the signature for LPs after grouping by inferred risk management style, together with 95% confidence bounds obtained via the adaptive bootstrap procedure using $B = 10,000$ replications. Panel D draws the corresponding signatures over a split sample. In Panels A – C, the dotted (solid) lines are the raw (smoothed) signature estimates. Source data provided by trader and Deutsche Bank.

erate decays of 40 – 50 cbps (LPs 3 – 6), and then LP7 with a uniquely strong decay of around 100 cbps. It resembles the classification of LPs made by Butz and Oomen (2018) into passive internalisers, impatient internalisers, and ag-

Table 1: Monitoring of LP risk management style

	LP 1	LP 2	LP 3	LP 4	LP 5	LP 6	LP 7
LP 1		≈	≠	≠	≠	≠	≠
LP 2	54.8%		≠	≠	≠	≠	≠
LP 3	0.0%	0.0%		≈	≈	≈	≠
LP 4	0.0%	0.1%	95.6%		≈	≈	≠
LP 5	0.0%	0.0%	8.0%	9.7%		≈	≠
LP 6	0.0%	0.0%	21.3%	24.3%	76.0%		≠
LP 7	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

Note. This table reports the L2 test in Eq. (13) to assess equality between post-deal micro signatures over a 30-second horizon, applied pairwise by LP. The lower diagonal lists the p-values of the test, obtained by the adaptive block bootstrap. The upper diagonal indicates the pair-wise signatures that are statistically indistinguishable (≈) and those that are different (≠) at a confidence level of 5%. The grey blocks ■ indicate groupings of LPs with statistically indistinguishable signatures. Source data provided by trader and Deutsche Bank.

Table 2: Consistency of LP risk management style

		1st half of sample			2nd half of sample		
		LP 1-2	LP 3-6	LP 7	LP 1-2	LP 3-6	LP 7
1st half	LP 1-2		≠	≠	≈	≠	≠
	LP 3-6	0.2%		≠	≠	≈	≠
	LP 7	0.0%	0.0%		≠	≠	≈
2nd half	LP 1-2	83.7%	0.5%	0.0%		≠	≠
	LP 3-6	0.0%	44.0%	0.0%	0.0%		≠
	LP 7	0.0%	0.0%	73.6%	0.0%	0.0%	

Note. This table reports the L2 test in Eq. (13) to assess equality between post-deal micro signatures over a 30-second horizon, applied pairwise by LP groupings and across a split-sample. The lower diagonal lists the p-values of the test, obtained by the adaptive block bootstrap. The upper diagonal indicates the pair-wise signatures that are statistically indistinguishable (≈) and those that are different (≠) at a confidence level of 5%. The grey blocks ■ indicate signatures that are statistically indistinguishable between LP groups and split-sample periods. Source data provided by trader and Deutsche Bank.

gressive internalisers or externalisers respectively. The FDA results indicate that differences within these groups are insignificant but differences between these groups are highly significant. This finding is reinforced by the grouped signatures with bootstrapped confidence bounds reported in Panel C of Figure 12.

To judge stability of the results, and the LP classification in particular, I perform a split sample analysis on the grouped signatures, see Table 2 and Panel D of Figure 12. The pairwise FDA tests suggest that the group signatures in the first and second half of the sample are statistically indistinguishable from one another. Moreover, differences between groups remain highly significant, not just contemporaneously within the same sub-sample but also across sub-samples (e.g. the signature of the passive internalisers in first half of the sample is significantly different from that of the impatient internalisers in second half of the sample).

A trader's ability to distinguish between competing liquidity offerings – not just in terms of spread or amount shown, but also by the LPs' footprint in the market when they risk manage the trader's flow – is a pre-requisite to efficient execution. This case study shows that the techniques presented in this paper allow the trader to classify LPs by their risk management style in a data-driven manner, and can then continue to monitor for any statistically significant changes to this classification over time. This in turn, enables the trader to selectively engage only with those LPs whose risk management style are compatible with their execution objectives.

Table 3: Detection of externaliser LP

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
Week 1		≈	≈	≈	≈	≈	≈	≠	≠
Week 2	65.3%		≈	≈	≈	≈	≈	≠	≠
Week 3	37.9%	23.0%		≈	≈	≈	≈	≠	≠
Week 4	86.6%	93.9%	27.8%		≈	≈	≈	≠	≠
Week 5	46.8%	28.5%	77.3%	30.7%		≈	≈	≠	≠
Week 6	94.1%	59.3%	68.4%	74.2%	85.8%		≈	≠	≠
Week 7	46.0%	21.8%	69.8%	29.4%	67.7%	78.6%		≠	≠
Week 8	0.2%	0.3%	0.1%	0.3%	0.1%	0.2%	0.1%		≈
Week 9	0.1%	2.5%	0.2%	2.0%	0.3%	1.4%	0.5%	18.0%	

Note. This table reports the L2 test in Eq. (13) to assess equality between post-deal micro signatures over a 30-second horizon, applied pairwise by weekly sample periods. The lower diagonal lists the p-values of the test, obtained by the adaptive block bootstrap. The upper diagonal indicates the pair-wise signatures that are statistically indistinguishable (≈) and those that are different (≠) at a confidence level of 5%. The grey blocks ■ indicate signatures that are statistically indistinguishable between weeks. Source data provided by Deutsche Bank.

3.3 Detection of externaliser LP

In the final case study, a trader conducts the following experiment. For seven weeks they leave the composition of their aggregator unchanged: it is composed of multiple largely internalising LPs. Then, for the subsequent two weeks they add a candidate externalising LP. The trader's execution style and trading activity is unaltered over this period, and the LPs are not made aware of the changes to the aggregator composition until the end of the experiment. The trader's objective is to understand whether the addition of the externalising LP has any measurable impact on execution performance.

Panels A and B of Figure 13 report the macro and micro signatures, the latter broken down by week. The macro signature exhibits a pronounced "take profit" like pre-deal pattern and then a sharp decay in the post-deal micro signature over a horizon of 30 seconds. The grouping of micro signatures by week reveals a distinct polarisation with only a modest decay for the signatures in the first seven weeks and then a much stronger decay for the two subsequent weeks when the externaliser is added into the aggregator. To assess whether these differences are statistically significant, I perform the same pairwise FDA analysis as in Section 3.2 but now on the nine weekly signatures. The results in Table 3 are very clear (see also Table 6 in Appendix B). The signatures for the first seven weeks are statis-

tically indistinguishable, and so are the signatures for week eight and nine, but the differences between those two groups is highly significant (see also Panel C of Figure 13).

What if the trader had not informed the LPs about the experiment, even ex-post? Or similarly, what if an LP switches their risk management approach from internalisation to externalisation but does not inform the trader? How quickly and how accurately could one detect such a change in trading behaviour? To answer this, I perform an FDA break-point test (see, e.g., Horváth and Kokoszka, 2012). It is calculated as follows:

$$T_{BP}(n) = \frac{n(N-n)}{N} \int \left(\underline{\mu}_n(\delta) - \overline{\mu}_n(\delta) \right) d\delta, \tag{24}$$

where

$$\underline{\mu}_n(\delta) = \frac{1}{n} \sum_{i=1}^{n} X_i(\delta) \qquad \text{and} \qquad \overline{\mu}_n(\delta) = \frac{1}{N-n} \sum_{i=n+1}^{N} X_i(\delta). \tag{25}$$

A structural break is then identified at the point where the test statistic is largest and statistically significant. The latter can be assessed via the same bootstrap procedures as described above. Panel D of Figure 13 draws the test value over the full sample period. There is a defined and unique maximum right at the start of week 8, the point where the externaliser was introduced into the aggregator. Closer inspection locates the breakpoint 83 trades and 25 minutes into the session. Considering that we analyse a 2 month period over which the trader executes more than 15,000 trades, the accuracy at which the test pinpoints the break is remarkable.

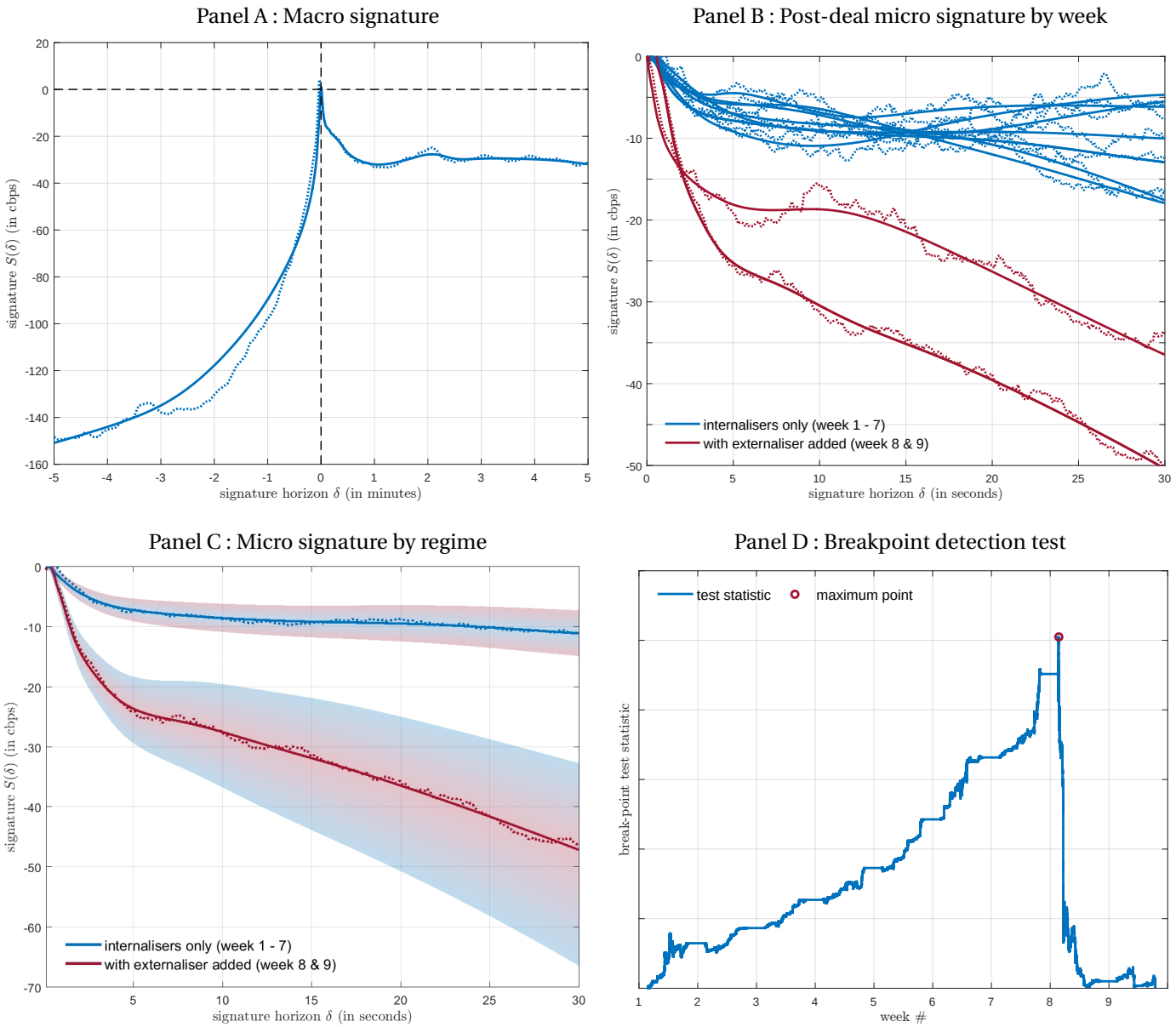
In Oomen (2017a) I show that simultaneously placing internalising and externalising LPs in competition can give rise to "prisoner's dilemma" effects and lead to the deterioration of liquidity access and increased execution costs. This case study shows how such a situation can be avoided in practice via a rigorous data driven process.

4 Conclusion

This paper demonstrates how functional data analysis and resampling techniques can be used to study signature profiles to undercover systematic dynamics in financial markets that are often deeply buried in vast amounts of data. The case studies illustrate the power of this approach in the context of electronic currency spot trading. I show how previously developed theory on optimal aggregator design can be deployed in practice via a rigorous data-driven approach that results in first order improvements in execution efficiency and liquidity access.

From a methodological perspective, there are various avenues for future research. For instance, it may be possible to extend the FDA distribution theory to allow for varying observation weights and make adjustments for the signature dependence. If successful, that would avoid the need for what can be a very time consuming bootstrap resampling procedure. Also, an implicit assumption I make is that s_n independent of q_n which may not hold in practice. From an empirical perspective, the case studies presented here are just one example of how the price signature

Figure 13: Detection of externaliser LP



Note. This chart draws the trader's macro signature pre- and post-deal over a five minute horizon in Panel A, and the weekly post-deal micro signature over the first 30 seconds in Panel B. Panel C draws the signatures classified by aggregator design scenario, with superimposed 95% confidence bounds obtained via the adaptive bootstrap procedure using $B = 10,000$ replications. Panel D draws the signature break-point test statistic in Eq. (24). In Panels A – C, the dotted (solid) lines are the raw (smoothed) signature estimates. Source data provided by Deutsche Bank.

analysis can be used in practice. But as already mentioned in the text, the methodology applies more broadly to a general class of signatures and facilitates detailed analysis of numerous other signature metrics of interest.

A Proofs

Proof of Propositions 1 and 2. The signature variance is a special case of the covariance function $\gamma(\delta_1,\delta_2)\equiv \text{cov}(S(\delta_1),S(\delta_2))$ derived below. Let $\underline{\delta}=\min(\delta_1,\delta_2)$, $\overline{\delta}=\max(\delta_1,\delta_2)$, $\underline{M}=\lfloor\underline{\delta}/\Delta\rfloor\wedge N$, $\overline{M}=\lfloor\overline{\delta}/\Delta\rfloor\wedge N$, $D=\lfloor(\overline{\delta}-\underline{\delta})/\Delta\rfloor\wedge N$.

$$
\begin{aligned}
\gamma(\delta_1,\delta_2) &= E(S(\delta_1)S(\delta_2)) = E(E(S(\delta_1)S(\delta_2)|\{d_n,q_n\}_{n=1}^N)),\\
&= E\left(\frac{\sigma^2}{(q'\iota)^2}\sum_n\sum_m d_n q_n d_m q_m |[t_n,t_n+\delta_1]\cap[t_m,t_m+\delta_2]|\right),\\
&= \frac{\sigma^2}{N^2}E(q'q)\min(\delta_1,\delta_2)+\frac{\sigma^2}{N^2}\sum_{n<m}E(d_nq_nd_mq_m)\max(\min(t_n+\delta_1-t_m,\delta_2),0)\\
&\quad+\frac{\sigma^2}{N^2}\sum_{n>m}E(d_nq_nd_mq_m)\max(\min(t_m+\delta_2-t_n,\delta_1),0)\\
&= \frac{\sigma^2}{N}\frac{\mu_q^2+\sigma_q^2}{\mu_q^2}\underline{\delta}+\frac{\sigma^2}{N^2}\sum_{n<m}(\sigma_q^2\rho_q^{|n-m|}+\mu_q^2)(\sigma_d^2\rho_d^{|n-m|}+\mu_d^2)\max(\underline{\delta}-(m-n)\Delta,0)\\
&\quad+\frac{\sigma^2}{N^2}\sum_{n>m}(\sigma_q^2\rho_q^{|n-m|}+\mu_q^2)(\sigma_d^2\rho_d^{|n-m|}+\mu_d^2)\max(\min(\overline{\delta}-(n-m)\Delta,\underline{\delta}),0)\\
&= \frac{\sigma^2}{N}\frac{\mu_q^2+\sigma_q^2}{\mu_q^2}\underline{\delta}+\frac{\sigma^2}{N^2}\sum_{k=1}^{M}(N-k)(\sigma_q^2\rho_q^k+\mu_q^2)(\sigma_d^2\rho_d^k+\mu_d^2)(\underline{\delta}-k\Delta)\\
&\quad+\frac{\sigma^2}{N^2}\sum_{k=1}^{D}(N-k)(\sigma_q^2\rho_q^k+\mu_q^2)(\sigma_d^2\rho_d^k+\mu_d^2)\underline{\delta}+\frac{\sigma^2}{N^2}\sum_{k=D+1}^{\overline{M}}(N-k)(\sigma_q^2\rho_q^k+\mu_q^2)(\sigma_d^2\rho_d^k+\mu_d^2)(\overline{\delta}-k\Delta)\\
&= \frac{\sigma^2}{N}\frac{\mu_q^2+\sigma_q^2}{\mu_q^2}\underline{\delta}+\frac{\sigma^2}{N^2}\left(\psi_1(\underline{M},\underline{\delta})+\underline{\delta}\xi_1(D)+\psi_1(\overline{M},\overline{\delta})-\psi_1(D,\overline{\delta})\right)\mu_d^2\\
&\quad+\frac{\sigma^2}{N^2}\left(\psi_{\rho_d}(\underline{M},\underline{\delta})+\underline{\delta}\xi_{\rho_d}(D)+\psi_{\rho_d}(\overline{M},\overline{\delta})-\psi_{\rho_d}(D,\overline{\delta})\right)\sigma_d^2\\
&\quad+\frac{\sigma^2}{N^2}\frac{\sigma_q^2}{\mu_q^2}\left(\psi_{\rho_q}(\underline{M},\underline{\delta})+\underline{\delta}\xi_{\rho_q}(D)+\psi_{\rho_q}(\overline{M},\overline{\delta})-\psi_{\rho_q}(D,\overline{\delta})\right)\mu_d^2\\
&\quad+\frac{\sigma^2}{N^2}\frac{\sigma_q^2}{\mu_q^2}\left(\psi_{\rho_q\rho_d}(\underline{M},\underline{\delta})+\underline{\delta}\xi_{\rho_q\rho_d}(D)+\psi_{\rho_q\rho_d}(\overline{M},\overline{\delta})-\psi_{\rho_q\rho_d}(D,\overline{\delta})\right)\sigma_d^2\\
&= \frac{\sigma^2}{N}\frac{\mu_q^2+\sigma_q^2}{\mu_q^2}\underline{\delta}+\frac{\sigma^2}{N^2}\left(\mu_d^2\phi_1(\delta_1,\delta_2)+\sigma_d^2\phi_{\rho_d}(\delta_1,\delta_2)\right)+\frac{\sigma^2}{N^2}\frac{\sigma_q^2}{\mu_q^2}\left(\mu_d^2\phi_{\rho_q}(\delta_1,\delta_2)+\sigma_d^2\phi_{\rho_q\rho_d}(\delta_1,\delta_2)\right)
\end{aligned}
$$

where $|I|$ denotes the length of interval I, $\phi_\rho(\delta_1,\delta_2)=\psi_\rho(\underline{M},\underline{\delta})+\underline{\delta}\xi_\rho(D)+\psi_\rho(\overline{M},\overline{\delta})-\psi_\rho(D,\overline{\delta})$ and,

$$
\xi_\rho(K) \equiv \sum_{k=1}^{K}\rho^k(N-k)=\begin{cases}\frac{1}{2}K(2N-K-1) & \rho=1\\ \rho\left(\left(1-\rho^K\right)\frac{N(1-\rho)-1}{(\rho-1)^2}+\frac{K\rho^K}{1-\rho}\right) & \rho\neq 1\end{cases}, \tag{26}
$$

$$
\psi_\rho(K,\delta) \equiv \sum_{k=1}^{K}\rho^k(N-k)(\delta-k\Delta)=\begin{cases}\frac{1}{2}NK(2\delta-\Delta(K+1))+\frac{1}{6}K(K+1)(2K\Delta+\Delta-3\delta) & \rho=1\\ \rho(1-\rho^K)\left(\frac{N\delta}{1-\rho}-\frac{\delta+N\Delta}{(1-\rho)^2}-\frac{(\rho+1)\Delta}{(\rho-1)^3}\right)+K\rho^{K+1}\left(\frac{\Delta(N-K)+\delta}{1-\rho}-\frac{2\Delta}{(1-\rho)^2}\right) & \rho\neq 1\end{cases}. \tag{27}
$$

■

Proof of Markov process autocorrelation function. For the two-state Markov process in (4), using the eigenvalue decompo-

sition (i.e. $P = U\Lambda U^{-1}$ and $P^k = U\Lambda^k U^{-1}$), the $k-$step transition probability matrix takes the form:

$$\begin{aligned} P^k &= \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}^k = \frac{1}{\alpha+\beta}\begin{pmatrix} \alpha & \alpha \\ \alpha & -\beta \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1-\alpha-\beta \end{pmatrix}^k\begin{pmatrix} \beta/\alpha & 1 \\ 1 & -1 \end{pmatrix}, \\ &= \frac{1}{\alpha+\beta}\begin{pmatrix} \beta+\alpha(1-\alpha-\beta)^k & \alpha-\alpha(1-\alpha-\beta)^k \\ \beta-\beta(1-\alpha-\beta)^k & \alpha+\beta(1-\alpha-\beta)^k \end{pmatrix}. \end{aligned} \tag{28}$$

From Eq. (28) and $\Pr(d_n = 1) = 1 - \Pr(d_n = -1) = \frac{\beta}{\alpha+\beta}$ it then follows that:

$$\begin{aligned} E(d_n d_{n+k}) &= E(d_n d_{n+k}|d_n = 1)\Pr(d_n = 1) + E(d_n d_{n+k}|d_n = -1)\Pr(d_n = -1), \\ &= E(d_{n+k}|d_n = 1)\frac{\beta}{\alpha+\beta} - E(d_{n+k}|d_n = -1)\frac{\alpha}{\alpha+\beta}, \\ &= \big(\beta-\alpha+2\alpha(1-\alpha-\beta)^k\big)\frac{\beta}{(\alpha+\beta)^2} - \big(\beta-\alpha-2\beta(1-\alpha-\beta)^k\big)\frac{\alpha}{(\alpha+\beta)^2}, \\ &= \frac{(\alpha-\beta)^2}{(\alpha+\beta)^2} + 4\frac{\alpha\beta}{(\alpha+\beta)^2}(1-\alpha-\beta)^k. \end{aligned} \tag{29}$$

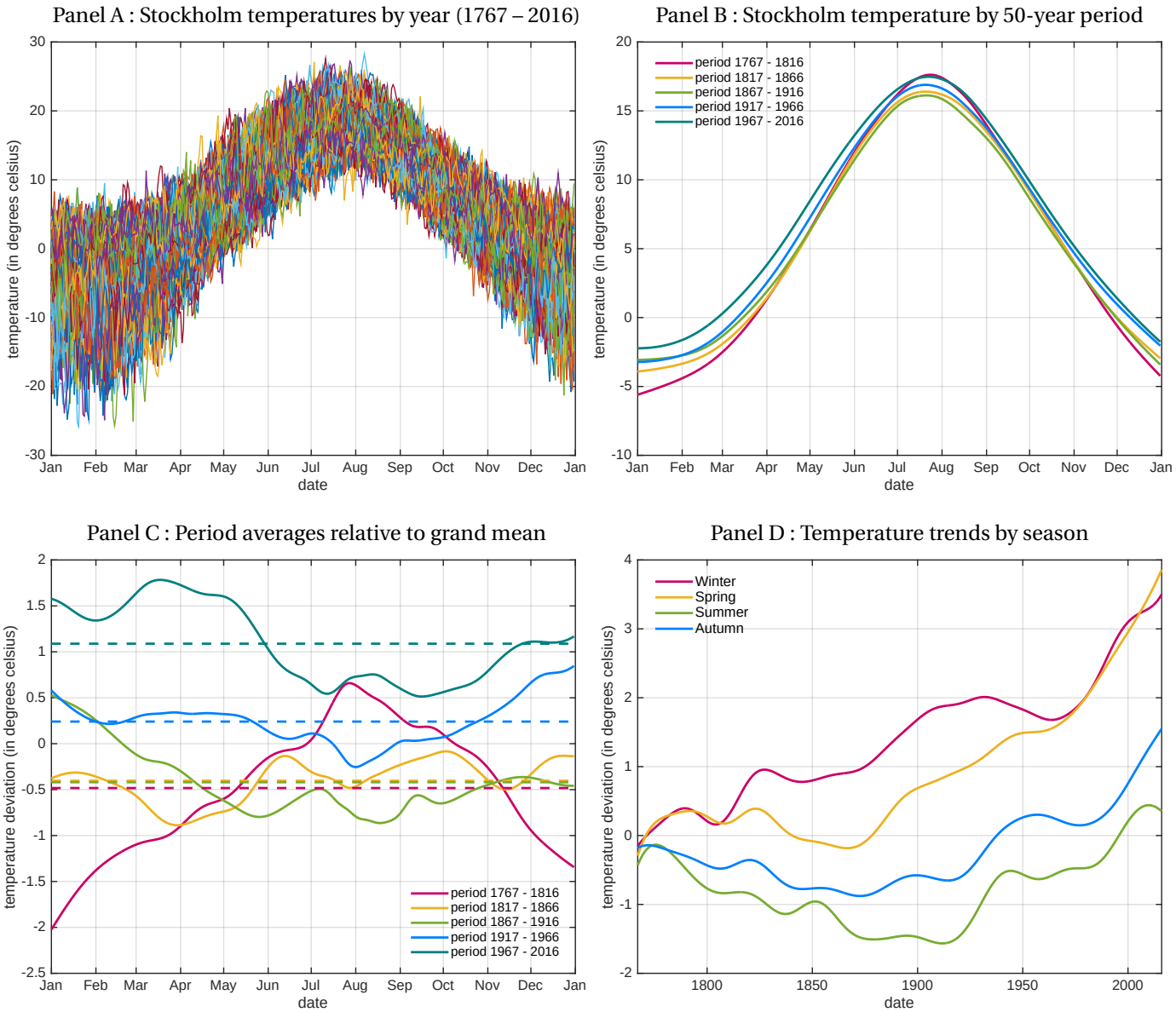
Finally, using the result in Eq. (29), $E(d_n) = \frac{\beta-\alpha}{\alpha+\beta}$, $V(d_n) = 4\alpha\beta/(\alpha+\beta)^2$, and for the choice of α and β as specified in Eq. (5), it follows that:

$$\mathrm{corr}(d_n, d_{n+k}) = \frac{E(d_n d_{n+k}) - E(d_n)^2}{V(d_n)} = (1-\alpha-\beta)^k = \rho_d^k. \tag{30}$$

■

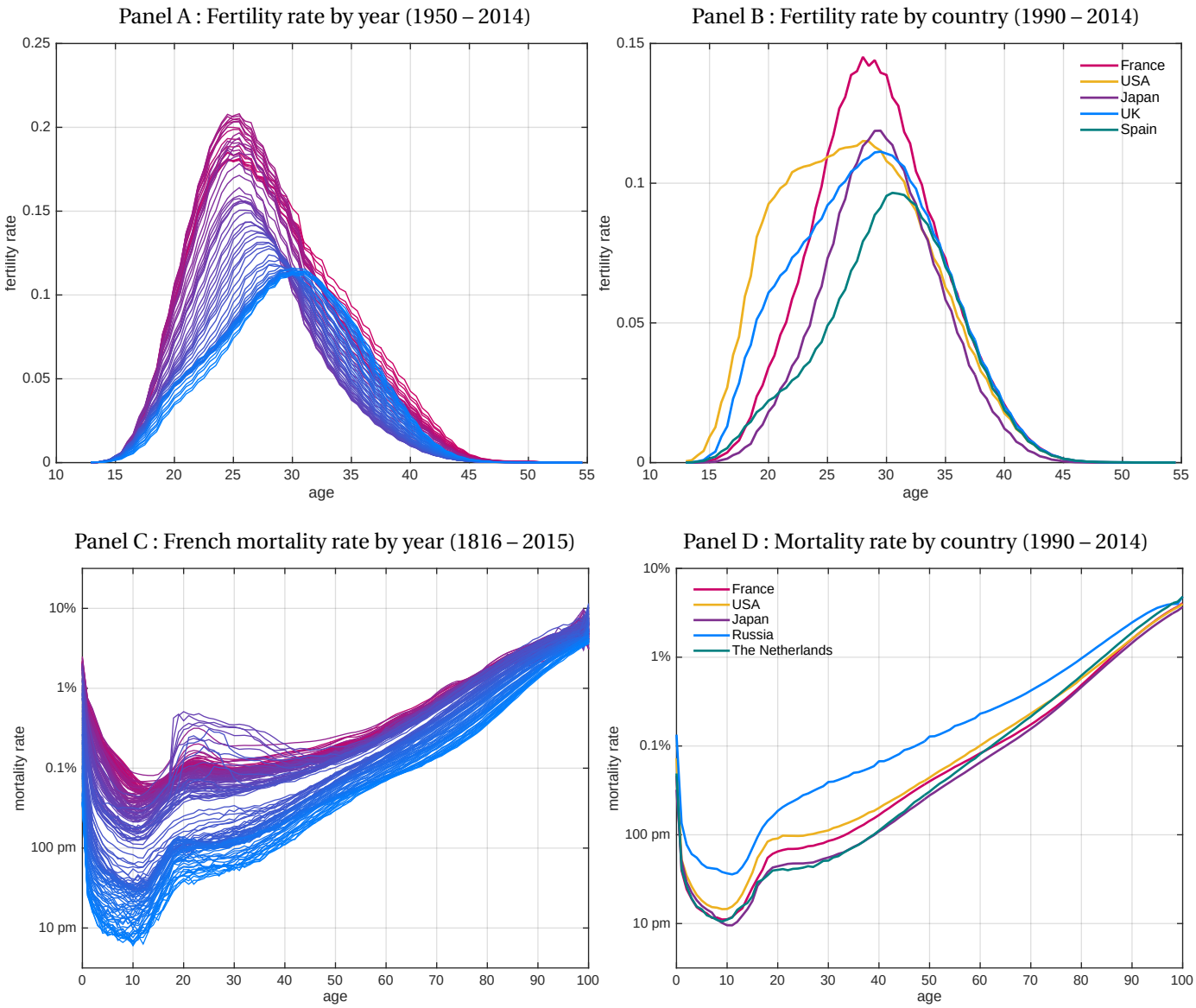
B Additional tables and charts

Figure 14: Illustrations of functional data – global warming



Note. This chart draws the daily Stockholm temperature measurements from Jan 1767 – Dec 2016. The average daily temperature by 50-year intervals and their differences to the full-sample grand mean is reported in Panels B and C respectively. Panel D reports the change average temperature over time, normalised at the start of the sample, and categorised by season. The source data is available from https://bolin.su.se/data/stockholm.

Figure 15: Illustrations of functional data – fertility & mortality rates



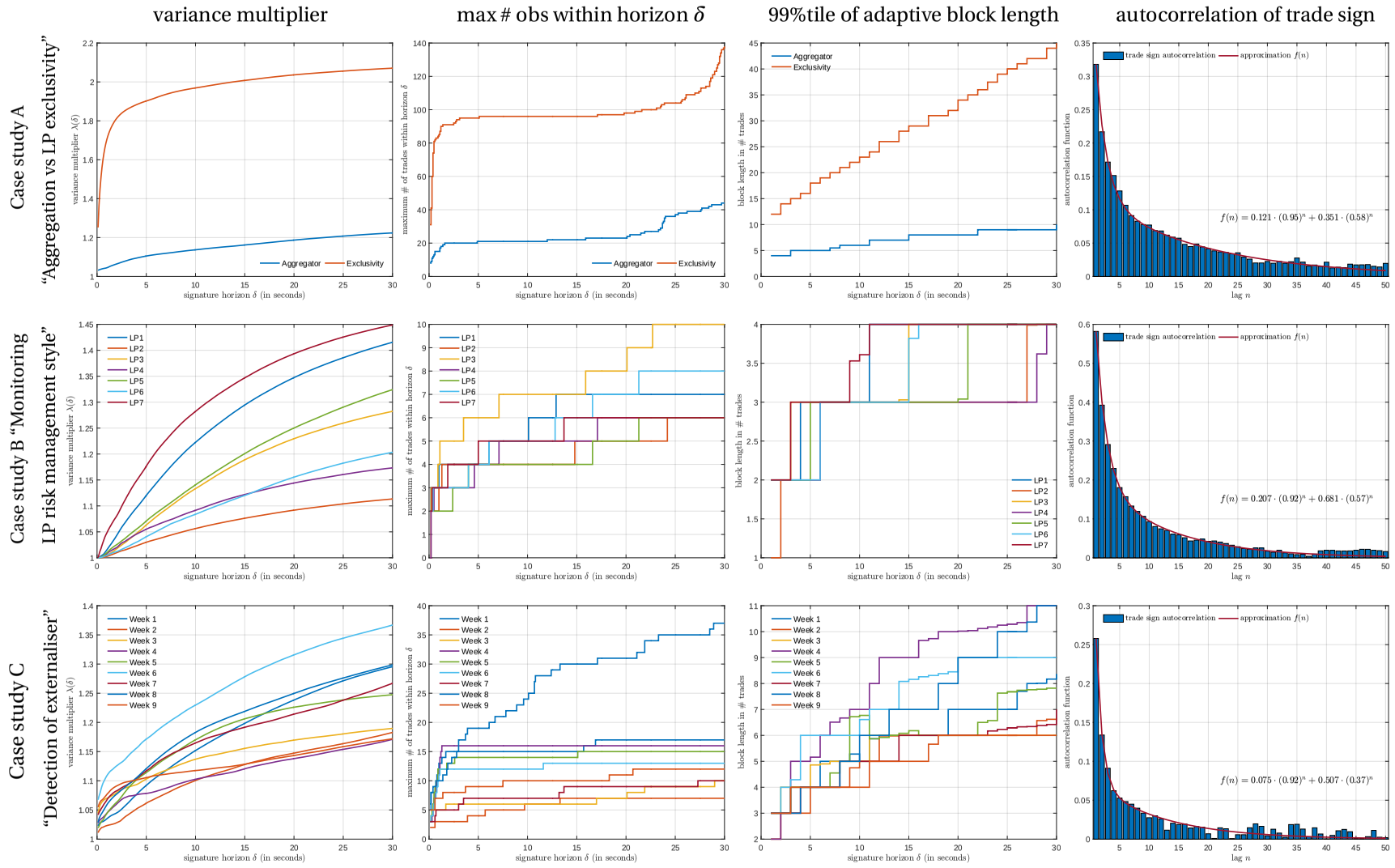
Note. Panels A and C draw the fertility rate and mortality rate respectively as a function of age and categorised by year, with the start of the sample colour coded in red and the end of the sample blue, and linearly interpolated in between. The fertility rate in Panel A is taken across multiple countries, and then broken out and averaged across the last 25 years of the sample in Panel B. The average mortality rate over the same period and categorised by country is reported in Panel D. The source data is available from http://www.humanfertility.org and http://www.mortality.org.

Table 4: Case study data summary statistics

			trade size								adaptive block bootstrap at 30s: block length by # trades				max
	# trades	$ volume	median	95%	max	$\mathcal{I}^{50\%}_{30s}$	$\mathcal{I}^{95\%}_{30s}$	λ^{30s}_{t}	λ^{30s}_{tdq}	BL*	avg	95%	99%	max	MM:SS
Case study A : Aggregation versus LP exclusivity															
Aggregator	11,006	773	20,000	300,000	1,000,000	–29.2	–17.9	2.5	1.2	70	1.8	5	10	52	02:44
Exclusivity	73,472	3,456	13,000	200,000	850,000	–4.7	–2.6	10.5	2.1	298	5.0	17	45	888	28:21
Total	84,478	4,229	15,000	200,000	1,000,000	–7.9	–5.5	9.5	1.8	343	4.1	14	38	888	28:21
Case study B : Monitoring of LP risk management style															
LP 1	6,913	7,287	1,000,000	2,000,000	5,000,000	–23.0	–19.4	1.4	1.4	22	1.3	3	4	11	01:34
LP 2	9,706	13,072	1,000,000	5,000,000	10,000,000	–24.8	–21.3	1.3	1.1	23	1.2	2	4	9	01:23
LP 3	7,073	14,098	1,000,000	10,000,000	20,000,000	–37.0	–31.2	1.4	1.3	33	1.3	3	4	9	01:16
LP 4	4,192	5,015	1,000,000	2,000,000	10,000,000	–37.1	–30.9	1.3	1.2	16	1.2	2	4	6	01:08
LP 5	2,348	3,306	1,000,000	3,000,000	5,000,000	–44.8	–37.5	1.3	1.3	13	1.2	3	4	10	01:23
LP 6	5,545	7,636	1,000,000	3,000,000	20,000,000	–42.5	–35.6	1.3	1.2	26	1.2	2	4	9	01:08
LP 7	1,097	1,960	1,000,000	5,000,000	5,000,000	–79.5	–66.7	1.5	1.4	11	1.3	3	4	6	00:51
Total	34,874	52,374	1,000,000	5,000,000	20,000,000	–33.8	–29.5	4.4	4.5	161	2.3	6	13	49	01:42
Case study C : Detection of externaliser															
Week 1	2,656	1471	500,000	1,099,253	3,000,000	–9.4	–4.1	2.1	1.3	23	1.6	4	8	26	02:54
Week 2	813	436	500,000	1,092,594	2,794,398	–9.6	–3.0	1.6	1.2	11	1.4	3	7	8	01:11
Week 3	821	362	367,632	1,093,800	2,000,000	–5.8	–0.7	1.9	1.2	9	1.4	3	6	13	00:52
Week 4	1,380	587	337,000	1,111,605	2,000,000	–9.7	–3.5	2.6	1.2	28	1.7	5	11	23	01:42
Week 5	1,157	555	369,046	1,097,569	3,000,000	–7.0	–2.0	2.2	1.2	16	1.5	3	8	32	02:01
Week 6	1,391	666	401,500	1,102,978	2,000,000	–8.3	–0.6	2.4	1.4	18	1.6	4	9	17	02:02
Week 7	1,186	611	500,000	1,061,800	2,000,000	–7.6	–1.9	1.9	1.3	11	1.5	4	7	21	02:10
Week 8	4,904	2,949	500,000	1,325,400	3,460,617	–33.3	–22.1	2.8	1.3	21	2.1	6	11	123	04:23
Week 9	1,671	906	460,623	1,000,000	3,360,357	–23.1	–12.8	1.9	1.2	12	1.6	4	6	17	01:41
Total	15,979	8,543	500,000	1,104,312	3,460,617	–17.7	–13.6	2.3	1.3	70	1.7	4	9	123	04:23

Note. This table reports summary statistics for the three case studies. For the trade data used in the signature construction, the table reports number of observations, total $ notional traded (in millions), and the median, 95th percentile, and maximum of trade sizes. Regarding signature metrics, it reports the integrated median and 95th percentile of the post-deal micro signature over 30 second horizon (i.e. $\mathcal{I}^{50\%}_{30s}$ and $\mathcal{I}^{95\%}_{30s}$, see Eq. (20)), and the signature dependence function λ_{30s} over the same horizon. For the bootstrap procedures, the table reports the Patton, Politis, and White (2009) optimal block length for the stationary bootstrap procedure (BL*), and for the adaptive bootstrap it reports the average, 95th, 99th, and maximum block size measured in number of trades and the maximum adaptive block size measured in calendar time (HH:SS). Source data provided by trader and Deutsche Bank.

Figure 16: Case study data summary statistics



Note. This chart draws in rows for each case study, the signature dependence function $\lambda(\delta)$ in the first column, the maximum number of observations within the signature horizon δ in the second column, the 99th percentile of the adaptive block sizes in the third column, and the autocorrelation function of the trade sign d_n with superimposed exponential fit in the last column. Source data provided by trader and Deutsche Bank.

Table 5: Monitoring of LP risk management style

	LP 1	LP 2	LP 3	LP 4	LP 5	LP 6	LP 7
LP 1		36.1%	0.0%	0.0%	0.0%	0.0%	0.0%
LP 2	54.9%		0.0%	0.0%	0.0%	0.0%	0.0%
LP 3	0.0%	0.0%		69.8%	5.5%	20.0%	0.0%
LP 4	0.0%	0.1%	95.5%		10.6%	28.1%	0.0%
LP 5	0.0%	0.0%	7.9%	9.7%		72.8%	0.0%
LP 6	0.0%	0.0%	21.3%	24.5%	76.4%		0.0%
LP 7	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	

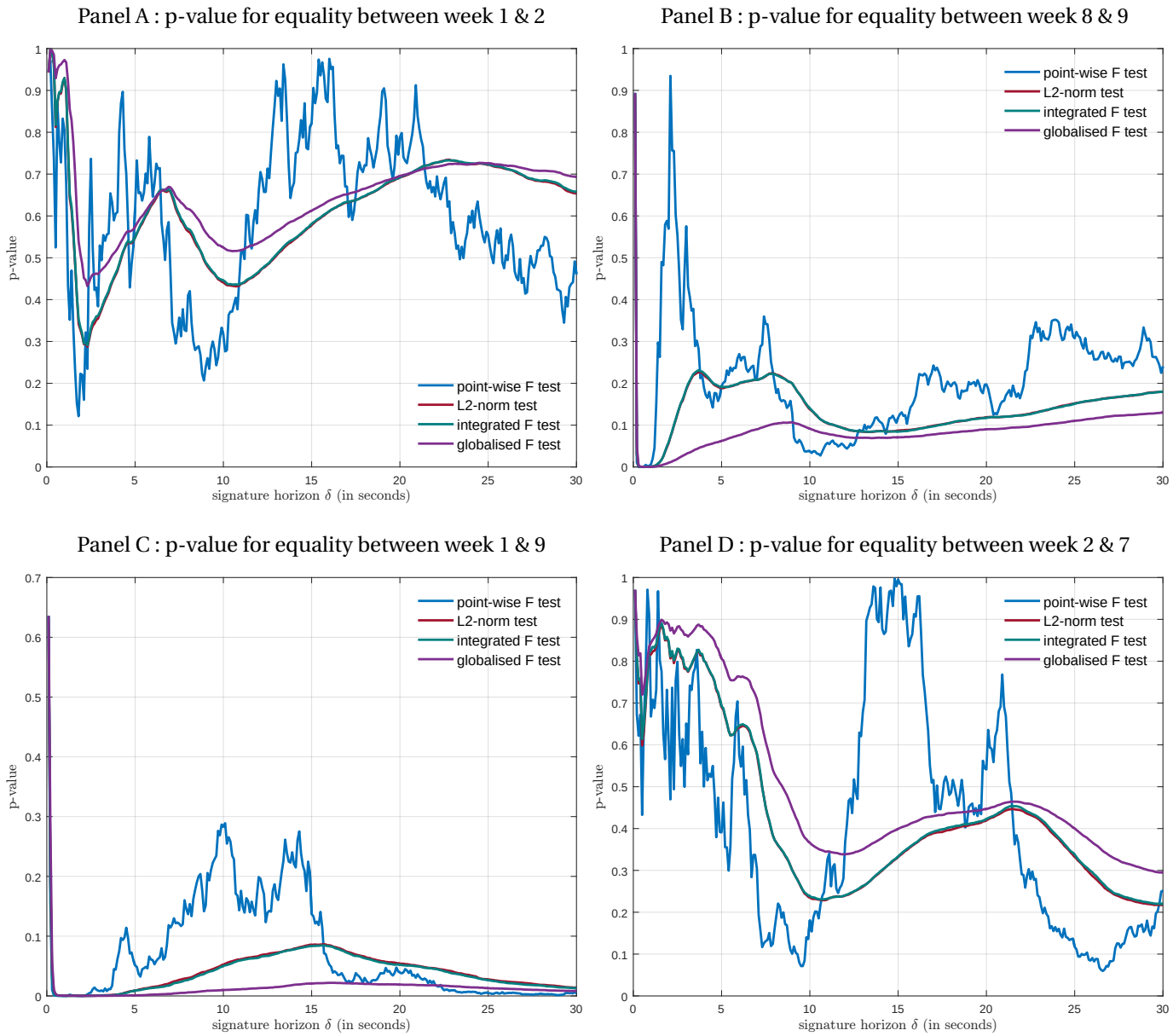
Note. This table reports the p-values of the integrated F test in the lower-diagonal and the globalised F test in the upper-diagonal as given by Eqs. (11) and (14) respectively, to assess equality between post-deal micro signatures over a 30-second horizon, applied pairwise by LP. The grey blocks ■ indicate groupings of LPs with statistically indistinguishable signatures at a confidence level of 5%. Source data provided by trader and Deutsche Bank.

Table 6: Detection of externaliser LP

	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
Week 1		69.3%	38.9%	92.8%	54.9%	90.3%	62.9%	0.1%	0.6%
Week 2	65.8%		37.2%	94.4%	44.1%	71.9%	29.5%	0.1%	0.8%
Week 3	38.1%	24.8%		40.2%	84.1%	69.7%	46.9%	0.0%	0.1%
Week 4	86.6%	94.0%	29.1%		46.9%	87.8%	44.6%	0.1%	0.8%
Week 5	47.2%	29.3%	78.3%	31.4%		88.4%	60.4%	0.0%	0.1%
Week 6	94.5%	60.4%	69.9%	74.7%	86.8%		87.8%	0.1%	0.8%
Week 7	46.6%	22.0%	70.7%	29.8%	68.5%	79.1%		0.1%	0.4%
Week 8	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%	0.1%		13.0%
Week 9	1.1%	2.4%	0.2%	1.9%	0.2%	1.3%	0.4%	18.0%	

Note. This table reports the p-values of the integrated F test in the lower-diagonal and the globalised F test in the upper-diagonal as given by Eqs. (11) and (14) respectively, to assess equality between post-deal micro signatures over a 30-second horizon, applied pairwise by weekly sample periods. The grey blocks ■ indicate signatures that are statistically indistinguishable between weeks at a confidence level of 5%. Source data provided by Deutsche Bank.

Figure 17: Comparison of alternative FDA tests for case study "Detection of externaliser LP"



Note. This chart draws the p-values of the FDA tests applied to pairwise weekly post-deal micro signatures as a function of signature horizon, i.e. the point-wise F test in Eq. (9), the integrated F test in Eq. (11), the L2-norm test in Eq. (13), and the globalised F test in Eq. (14). Source data provided by Deutsche Bank.

SALES AND TRADING DEPARTMENT – DISCLAIMER

This document is intended for discussion purposes only and does not create any legally binding obligations on the part of Deutsche Bank AG and/or its affiliates ("DB"). Without limitation, this document does not constitute an offer, an invitation to offer or a recommendation to enter into any transaction. DB is not acting as your legal, financial, tax or accounting adviser or in any other fiduciary capacity with respect to any proposed transaction mentioned herein. This document does not constitute the provision of investment advice and is not intended to do so, but is intended to be general information. Any product(s) or proposed transaction(s) mentioned herein may not be appropriate for all investors and before entering into any transaction you should take steps to ensure that you fully understand the transaction and have made an independent assessment of the appropriateness of the transaction in the light of your own objectives, needs and circumstances, including the possible risks and benefits of entering into such transaction. You should also consider seeking advice from your own advisers in making any assessment on the basis of this document. If you decide to enter into a transaction with DB, you do so in reliance on your own judgment. The information contained in this document is based on material we believe to be reliable; however, we do not represent that it is accurate, current, complete, or error free. Assumptions, estimates and opinions contained in this document constitute our judgment as of the date of the document and are subject to change without notice.

This material was prepared by a Sales or Trading function within DB, and was not produced, reviewed or edited by the Research Department (which is independent from the Sales or Trading function). Any opinions expressed herein may differ from the opinions expressed by other DB departments including the Research Department. Sales and Trading functions are subject to additional potential conflicts of interest which the Research Department does not face. DB may engage in transactions in a manner inconsistent with the views discussed herein. In general, Sales and Trading personnel are compensated in part based on the volume of transactions effected by them. DB seeks to transact business on an arm's length basis with sophisticated investors capable of independently evaluating the merits and risks of each transaction, with investors who make their own decision regarding those transactions.

The distribution of this document and availability of these products and services in certain jurisdictions may be restricted by law. You may not distribute this document, in whole or in part, without our express written permission. DB SPECIFICALLY DISCLAIMS ALL LIABILITY FOR ANY DIRECT, INDIRECT, CONSEQUENTIAL OR OTHER LOSSES OR DAMAGES INCLUDING, WITHOUT LIMITATION, LOSS OF PROFITS INCURRED BY YOU OR ANY THIRD PARTY THAT MAY ARISE FROM, OR IN CONNECTION WITH, ANY RELIANCE ON THIS DOCUMENT OR FOR THE RELIABILITY, ACCURACY, COMPLETENESS OR TIMELINESS THEREOF. DB is authorized under German Banking Law (competent authority: the European Central Bank and the BaFin - Federal Financial Supervising Authority) and DB in the UK is authorised by the Prudential Regulation Authority and is subject to limited regulation by the Prudential Regulation Authority and Financial Conduct Authority. Details about the extent of our authorisation and regulation are available on request or from https://www.db.com/disclosures.

References

Butz, M., and R. C. Oomen, 2018, "Internalisation by electronic FX spot dealers," *Quantitative Finance*, forthcoming.

Efron, B., 1979, "Bootstrap Methods : Another look at the jackknife," *Annals of statistics*, 7 (1), 1 – 26.

European Parliament and the Council, 2014, "Directive 2014/65/EU of the European Parliament and of the Council of 15 May 2014 on markets in financial instruments and amending Directive 2002/92/EC and Directive 2011/61/EU," *Official Journal of the European Union*, available at http://eur-lex.europa.eu/eli/dir/2014/65/oj.

Extance, A., 2016, "Digital DNA," *Nature*, 537, 22 – 24.

Faraway, J., and Q. Shen, 2004, "An F test for linear models with functional responses," *Statistica Sinica*, 14 (4), 1239 – 1257.

Górecki, T., and L. Smaga, 2015, "A comparison of tests for the one-way ANOVA problem for functional data," *Computational Statistics*, 30, 987 – 1010.

Griffiths, L., C. Rich, M. Geraci, F. Sera, M. Cortina-Borja, T. Pouliou, L. Platt, J. Johnson, and C. Dezateux, 2013, "Technical report on the enhancement of Millennium Cohort Study data with accelerometer-derived measures of physical activity and sedentary behaviour in seven year olds," available at https://sp.ukdataservice.ac.uk/doc/7238/mrdoc/pdf/mcs4_pa_technical_report.pdf.

Horváth, L., and P. Kokoszka, 2012, *Inference for functional data with applications*. Springer, New York, NY.

Horváth, L., P. Kokoszka, and G. Rice, 2014, "Testing stationarity of functional time series," *Journal of Econometrics*, 179, 66 – 82.

Künsch, H. R., 1989, "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of statistics*, 17 (3), 1217 – 1241.

Oomen, R. C., 2017a, "Execution in an aggregator," *Quantitative Finance*, 17 (3), 383 – 404.

———, 2017b, "Last look," *Quantitative Finance*, 17 (7), 1057 – 1070.

Patton, A., D. N. Politis, and H. White, 2009, "Correction to "Automatic block-length selection for the dependent bootstrap" by D. Politis and H. White," *Econometric Reviews*, 28 (4), 372 – 375.

Politis, D. N., and J. P. Romano, 1992, "A circular block-resampling procedure for stationary data," in *Exploring the Limits of Bootstrap*, ed. by R. LePage, and L. Billard. John Wiley, New York, pp. 263 – 270.

———, 1994, "The stationary bootstrap," *Journal of the American Statistical Association*, 89, 1303 – 1313.

Politis, D. N., and H. White, 2004, "Automatic block-length selection for the dependent bootstrap," *Econometric Reviews*, 23 (1), 53 – 70.

Posedel, P., 2005, "Properties and estimation of GARCH(1,1) model," *Metodološki zvezki*, 2(2), 243 – 257.

Ramsay, J. O., and B. W. Silverman, 2002, *Applied functional data analysis : Methods and case studies*. Springer, New York, NY.

———, 2005, *Functional data analysis*. Springer, New York, NY.

Reinsel, D., J. Gantz, and J. Rydning, 2017, "Data Age 2025: The Evolution of Data to Life-Critical," IDC White Paper, available at https://www.seagate.com/files/www-content/our-story/trends/files/Seagate-WP-DataAge2025-March-2017.pdf.

Zhang, J. T., 2014, *Analysis of variance for functional data*. CRC Press, Boca Raton, FL.

Zhang, J. T., and J. Chen, 2007, "Statistical inferences for functional data," *Annals of statistics*, 35 (3), 1052 – 1079.

Zhang, J. T., and X. Liang, 2014, "One-way anova for functional data via globalizing the pointwise F-test," *Scandinavian Journal of Statistics*, 41 (1), 51 – 71.

Zhang, X., 2016, "White noise testing and model diagnostic checking for functional time series," *Journal of Econometrics*, 194, 76 – 95.