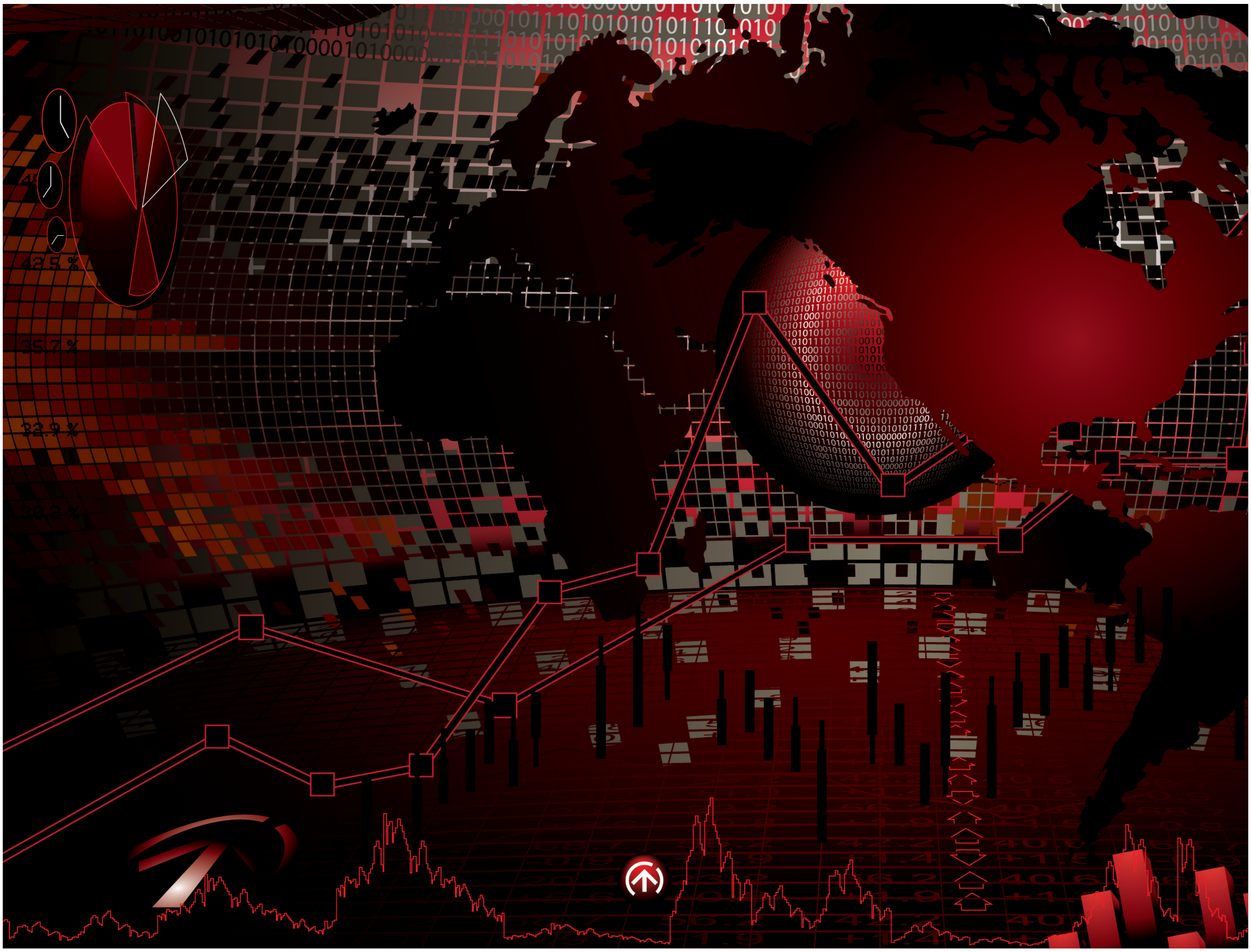
Risk publications

XVA

Special report 2018

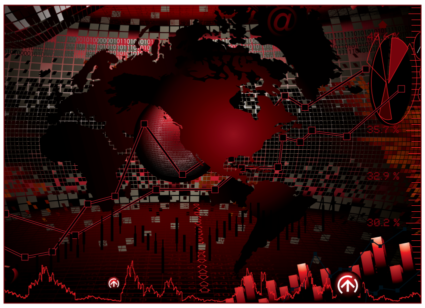
Risk.net



Sponsored by

Risk.net

ICA
The Independent Calculation Agent



Lukas Becker
Desk Editor, Derivatives
lukas.becker@infopro-digital.com

Joanna Edwards
Senior Director, Marketing Solutions
joanna.edwards@infopro-digital.com

Stuart Willes
Commercial Editorial Manager
stuart.willes@infopro-digital.com

Alex Hurrell, **Commercial Subeditor**
alex.hurrell@infopro-digital.com

Antony Chambers, **Publisher, *Risk.net***
antony.chambers@infopro-digital.com

Esha Khanna, **Commercial Manager**
esha.khanna@infopro-digital.com

Nadia Pittorru, **Sales Manager**
nadia.pittorru@infopro-digital.com

Celine Infeld, **Managing Director**
celine.infeld@infopro-digital.com

Ben Wood, **Group Publishing Director**
ben.wood@infopro-digital.com

Ben Cornish, **Senior Production Executive**
ben.cornish@infopro-digital.com

Infopro Digital (UK)
Haymarket House, 28–29 Haymarket
London SW1Y 4RX
Tel: +44 (0)20 7316 9000

Infopro Digital (US)
55 Broad Street, 22nd Floor
New York, NY 10004-2501
Tel: +1 646 736 1888

Infopro Digital (Asia-Pacific)
Unit 1704-05, Berkshire House
Taikoo Place, 25 Westlands Road,
Hong Kong, SAR China
Tel: +852 3411 4888

Cover image
liangpv/Getty

Published by Infopro Digital

Benchmark reform poses challenges for valuation teams

After a turbulent decade, the past couple of years have arguably been a little easier for the derivatives valuation community. Overnight indexed swap (OIS) discounting is market standard for cash collateralised instruments, while funding and capital valuation adjustments are no longer a matter of debate. Margin valuation adjustment could be on the horizon, but the Street is largely in 'wait-and-see' mode.

However, benchmark reform is a potential fly in the ointment. For euro swaps collateralised with cash, the rate used for discounting future cashflows – Eonia – will be banned for use in new trades from 2020, as it will not comply with the European Union's benchmarks regulation.

Lawyers point out the discount rate is not specified in contracts – it's up to the parties themselves to choose what rate to use – meaning swap users should be free to use Eonia after 2020. But, as there will be no new Eonia-referencing swaps traded, no curve can be constructed, making it impossible to continue to use for discounting.

A replacement rate has yet to be selected by the European Central Bank (ECB)-convened working group on euro risk-free rates, but sources close to the group believe it's highly likely to choose the euro short-term rate (Ester).

The problem is, this rate is under consultation and will not be ready until October 2019, according to the ECB. With Eonia barred from the start of 2020, the market will only have three months to build a functioning Ester curve that can be used not only for discounting, but for OIS trading in general.

The euro working group has already warned that the issue will create potential valuation problems from 2020, and it remains to be seen how the market will deal with this.

Another problem lies with the potential discontinuation of Libor. In July last year, the UK Financial Conduct Authority said it would give up its powers to compel panel banks to submit quotes to the range of Libor currencies from the end of 2021, which means the rate's survival will not be certain from that date.

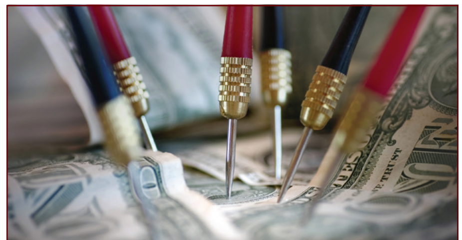
Libor-based curves are used to obtain estimates for the forward-floating rates, to determine future cashflows. This means there is a risk that Libor forecast curves may no longer be produced in their current form.

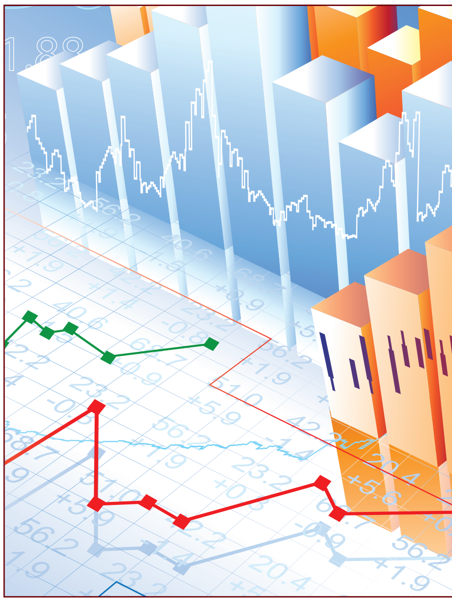
If Libor is discontinued after 2021, fallback clauses currently being developed by the industry should kick in, at which point existing contracts linked to the benchmark will reference the alternative risk-free rate, such as Sonia for the sterling market. A fixed spread will then be added to the rate to take into account bank credit risk.

The methodology for calculating the spread is still being developed but, when it is, a recent report by Numerix suggests using the fallback methodology to construct an alternative reference rate curve to help get a better view of future cashflows under the new rate. When that curve should kick in is currently anyone's guess, given some banks are still willing to keep Libor alive beyond 2021.

So, while benchmark reform is starting to impact parts of the derivatives business, it's unlikely valuations teams will be spared much longer.

Lukas Becker
Desk editor, derivatives, *Risk.net*





ICA
The Independent Calculation Agent

Technology at the service of XVA or ‘XVA as a service’?

Stéphane Rio, founder and chief executive officer of ICA, explains how new technologies can radically change approaches to XVA and other pricing and risk calculations when big data experts and quants work hand in hand

When facing the challenges of XVA computations with older-generation systems and limited in-house computational resources, banks have had to rethink their setups or cut corners completely. In particular, many were compelled to strip their overnight batches down to the bone and discard valuable intermediary results.

Big data technologies allow quants to think differently

Intermediate results are a means for quants and traders to analyse results and track down errors efficiently. They also avoid much of the recalculation when something minor is amended – a single trade or credit support annex term, for example. In a large netting set, this could mean less than 1% of the compute time. For this reason, the ICA system will save all future mark-to-markets of individual transactions and the sensitivities of all netting sets – on each path and on each date of the simulations. For large portfolios this means saving, slicing and dicing hundreds of billions of data points.

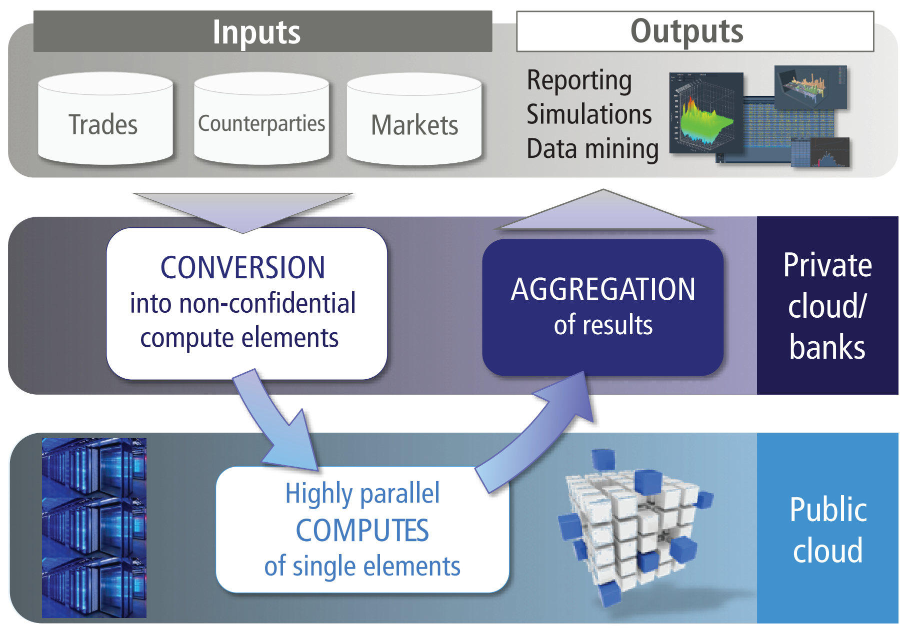
With big data technologies, this can now be achieved at a very marginal cost – both in terms of computational time and financial outlay. With a fast investigation tool and highly efficient runs, a world of new possibilities opens up: traders and salespeople can work on real-time pre-trade prices and focus on efficiently optimising XVAs and resources by, for example, selecting who to trade with to minimise margin valuation adjustment or simulating thousands of portfolio scenarios to optimise capital allocation.

Moving to the cloud – A game changer for agile teams

Banks using in-house computational power have finite resources. In their overnight batches, they will need to be frugal with computationally intensive calculations such as cross gammas, while most resources will remain underused intraday. Working with the cloud offers multiple benefits: elasticity, scalability, on-demand accessibility, redundancy and pay-per-use – one uses and pays for what one needs when it is needed. Two conditions are required for this.

First, to share two types of knowledge among a single team: models on the one hand, cloud and large-scale distribution technology on the other. This will allow quants to replace traditional grids, work much closer to models, optimise distribution (to avoid redundant calculations), minimise input/output and develop parallel orchestration with linear performance. Computation time can be reduced to the bare minimum; for example, using 10 times as many cores and reducing computation time tenfold will cost virtually the same amount.

Inputs and outputs of private and public cloud



Second, to meet regulators' requirements and ensure the security of data.

Guidelines from the US Federal Reserve, the UK Financial Conduct Authority and the European Central Bank are logical and clear on how to satisfactorily use public cloud-based technology.

As detailed in the February edition of *Risk*, ICA added an additional layer of security in ensuring no sensitive information is sent to a public cloud[1] – where the bulk of computations are performed, aggregations and final metrics computations being undertaken locally using ICA's big data database). ■

[1] *Sherif N, When a cloud can light the way,* Risk, *February 2018, www.risk.net/5395461*

BIG DATA AND CLOUD TECHNOLOGY SOLUTIONS

Big data and cloud technologies are two examples of very mature innovations that banks do not sufficiently leverage – and not only for XVA. ICA combines the skills of quants, new technology experts and experienced former front officers, and is therefore uniquely positioned to offer banks the full benefit of these technologies with two appropriate solutions:

XVA as a service

For users looking for a comprehensive and agile XVA solution, the ICA tool is available 'as a service': turn-key real-time XVA trading and risk solutions, including all value-added metrics, risk sensitivities, profit and loss attribution, *ad hoc* cross gamma, stress tests or other market scenarios, and real-time what-if scenarios (pre-trade pricing, assignments, terminations, change of credit support annex terms, central counterparty upload, and so on).

New technologies at the service of XVA

For banks wishing to continue using proprietary models while moving architecture to a new generation, ICA can inject some or all of those technologies into the bank's architecture through special implementation projects. ICA has distinctive experience of getting big data and cloud technologies to operate efficiently in a complex pricing and risk environment – and a successful track record.

Contact

Stéphane Rio
Founder and CEO

Emmanuel Danzin
Head of Business Development

E info@the-ica.com
W www.the-ica.com

Corporate benefits

The attractions of corporate trades have seen dealers flock to the market, but is the ensuing race to the bottom in swap prices sustainable? By Nazneen Sherif with editing by Lukas Becker

Need to know

- The corporate swaps market has seen a significant increase in competition in the past year, resulting in spreads recently tightening to levels some dealers say are unsustainable.
- Add-ons to swap prices to reflect credit risk, funding and return on capital have almost halved over the past year.
- The return on capital that banks price into swaps has also shrunk from double digits to single digits on certain trades.
- A number of market participants have been able to execute swaps close to mid, which some say aligns with pre-crisis levels.
- Some banks say corporates have been one of the few remaining areas where profits can be made, and where interesting axes can be obtained, leading to a stampede.
- Expected rate hikes could also give a positive funding profile to the trades.
- European banks say they have seen a particular push from US banks, who have a balance sheet and funding advantage versus European banks at present.
- Changing market practice has also played a role in the spread tightening – for instance, widespread use of hedge advisers and swap auctions, which increase competition.

One afternoon in early April, the treasurer of a large corporate was in a meeting with its hedge adviser. The corporate needed to borrow euros through a cross-currency swap and had asked a handful of banks for quotes.

One European bank eventually won the trade at a very narrow spread. Shortly after, the corporate was contacted by the bank with a new sales pitch, asking if they would be interested in buying corporate credit cards. The treasurer thought it was a bizarre request. But given that dealer competition for corporate trades has seen spreads on swaps halve in the past year alone, banks need to make their money back somehow.

"We have definitely seen some people bidding pretty suicidal levels," says one derivatives structurer at a US bank.

In many cases, the spreads have fallen well below banks' return-on-equity hurdle rates.

"I have seen banks where the logical pricing for them to get an adequate coverage of their capital would be at one level, and they will drop 10 basis points from that just to win the trade. This means, in some cases, they are pretty much making close to no return," says a treasurer of one European corporate.

The recent changes have reversed a trend that has seen the costs of corporate trades increase in recent years, driven by banks' efforts to cover rising counterparty credit, funding and capital cost hurdles.

So why are corporates suddenly the flavour of the month? Some see them as one of the few remaining client types where decent profits were available. Others say the trades give them unique axes that other dealers won't have.

There are more technical reasons too. Some say the prospect of rising rates in the US and Europe will deliver expected funding benefits from trades with these clients, increasing competition from all – including those banks that have wound down their non-core units and are looking to rebuild their uncollateralised corporate exposures.

European banks claim they've seen a particular push from their US rivals – Goldman Sachs has been vocal about its ambitions – which are using their large balance sheet and a funding cost advantage to drive down prices and win market share. "In the last couple of years, US banks have been able to increase market share not only in the US, but also in Europe and globally," says one senior fixed-income trader at a European bank. "And while they would say they just want to do more business with clients, it may be more to do with leveraging the dollar capacity and the dollar surplus they had."

Others say the increasing practice of using hedge advisers, and either increasing the number of banks invited to quote on a new swap, or splitting up the market and credit risk across dozens of banks, is pushing spreads down – to the extent that some banks simply can no longer be competitive.

Market participants talk of recently executed trades where bids were close to the mid-price and still lost, which some say is reminiscent of pre-crisis levels.

Tighter spreads, lower returns

Trades with corporates are typically uncollateralised, which caused spreads to widen significantly post-crisis as banks started to factor in additional counterparty credit risk, funding and capital costs associated with these trades into swap prices – in the form of credit valuation adjustment (CVA), funding valuation adjustment (FVA) and capital valuation adjustment (KVA), respectively. Collectively known as XVAs, these charges are typically passed onto corporate clients as part of the swap price.

The CVA capital charge for future variation in CVA was introduced in Basel III – except in Europe, where European banks are exempt from the charge for trades with local corporates – while FVA and KVA were commonly priced from 2012 and 2015, respectively. As a result, corporates faced rising spreads on uncollateralised hedges.

But that all changed around a year ago, when XVA charges began to crater.

"It's trading certainly at levels half of where we could get to," says the head of the corporates business at a second European bank. "We are seeing the more competitive pricing happening in the high-yield space, and that's where we are seeing the most dramatic change."

For instance, a five-year cross-currency swap with a high-yield name that used to clear at a spread of 50–60bp of XVAs a year ago is now trading at close to 30bp, says the rates head. On a swap where a high-yield corporate receives fixed against floating, the spread shrank from 3–5bp to 1–2bp.

Edouard Nguyen, former head of the dealing room at Paris-based corporate Veolia and now head of the corporates and treasuries practice at consulting firm Axis Alternatives, says the XVAs being charged to a corporate, mainly reflecting its counterparty credit risk and cost of funding to the bank, have shrunk by around 30% in a year.

"One year ago… XVAs for a 10-year interest rate swap fixed-rate receiver for a corporate like Veolia – rated BBB by S&P – ranged from 6.5–7bp running. More recent quotes show the prices have narrowed significantly. In the example of a BBB corporate, current quotes show XVAs around 4.5–5bp for the same swap," he adds.

The head of the corporates business at the second European bank says dealers have also been willing to take a lower required rate of return on the capital allocated to the trade just to win market share.

US banks have an "embedded competitive advantage", says a senior fixed-income trader

"When this first started out, I think everybody was sort of on the 15–20% hurdle. I think it has now come lower. What we are seeing is people are pricing single-digit returns on some of the trades," he adds.

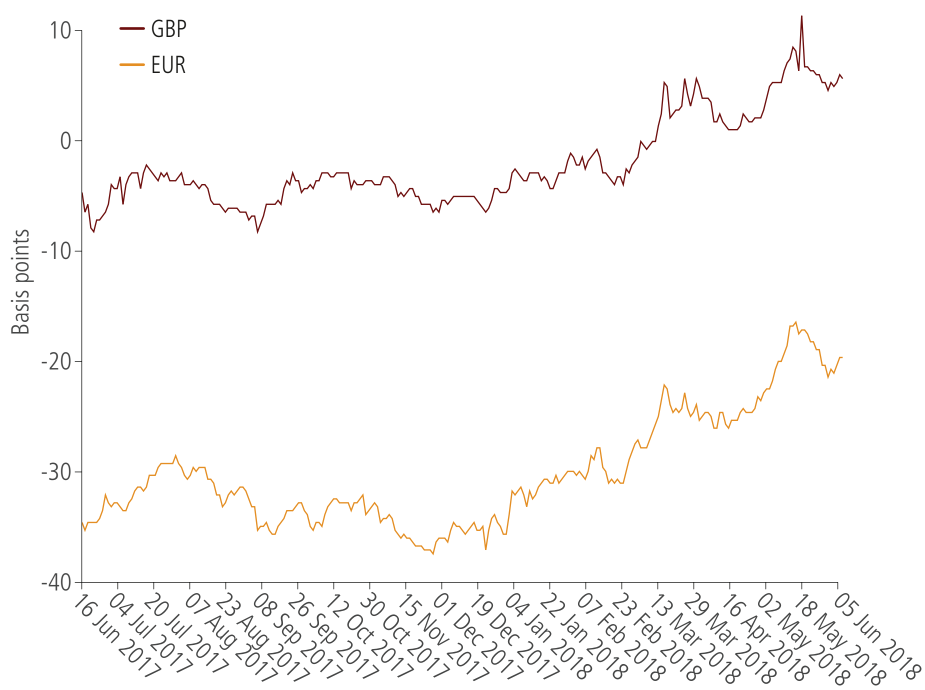
As a result, some banks are having to turn to cross-selling other services such as cash management to corporates to make up the difference (see box: If you don't cross, you make a loss).

"When you actively start cross selling it's a sign that profits are getting very slim," says François Jarrosson, a director in the hedging and derivatives team at Rothschild Global Advisory.

Back in vogue

The reasons why corporates are now so desirable are many and varied. A key one is that they can be lucrative clients to have. Speaking around 18 months ago, one senior rates structurer was

1 Five-year cross-currency basis spreads



Source: Bloomberg

bemoaning the lack of profits his bank was making from flow trading. When asked how they keep the lights on, he said: "corporates".

This wasn't a secret – a markets head at a European bank also says that with spreads getting thinner on flow trading, corporates were one of the few remaining areas where profits were available. The markets head also notes that the flows give interesting axes that they can then pass onto other market participants that other dealers would not have. As a result, dealers have piled in.

The US bank's derivatives structurer notes that banks that had to close down their non-core businesses post-crisis are now trying to claw their way back into a space they had vacated, in order to rebuild relationships.

"It feels like we are seeing more and more dealers coming back into this business with banks closing down their capital resolution or non-core divisions – that feels like a bit of an end of the austerity era," says the structurer. "And where in the past it has been difficult for them to participate, with that kind of capital resolution work stream having been wound up and finished, they are able to be back in the business of loading up their balance sheet again."

Others say an expectation of rising rates in the US may have also prompted some banks to price cheaply on trades with corporates that pay fixed and receive floating. This means the uncollateralised corporate could be out of the money more in its early years, which would see the bank not receiving any margin, but having to pay it on its collateralised hedge, creating a funding requirement.

"We have definitely seen some people bidding pretty suicidal levels"

Derivatives structurer at a US bank

But as rates rise, this flips around, creating a funding benefit. Plus, as the bank is effectively building a payable to the corporate during the later years of the trade, this means the credit risk from the corporate is lower.

"Clearly in the US [we are] in a rising rate environment, so maybe people are saying 'Oh, this is going to be in-the-money to the client and we are going to be money-good on it, so we will just go cheap," says the head of the corporates business at the second European bank.

Funding advantages

US banks are also said to be using their balance sheet advantages to muscle in on corporate business, helped by a collateral funding advantage from their long US dollar bias.

These banks use US dollar-euro cross-currency swaps to fund collateral to post as margin. For example, bank A lends dollars to bank B and receives euros, and vice versa. Bank A needs to make coupon payments in euros at the euro interest rate to the second bank, and receives the opposite. A basis, which is added to the euro interest rate paid by bank A, factors in the difference in funding costs of the two currencies.

On June 5, 2017, the euro cross-currency basis for a five-year trade against US dollar Libor was –35.25 and for sterling it was –9.25, according to data from Bloomberg. By April 12, 2018, the euro basis had fallen to –27.5, while sterling sat at –4.25 (figure 1).

This means US banks have been able to obtain euros and sterling at a rate better than the rate at which European banks can borrow dollars, making collateral payments cheaper to make compared to non-US banks.

"They have been enjoying the fact that on average they are long dollar and the dollar has been more complicated to find than any other currency in the past couple of years, which means they have an embedded competitive advantage compared to a lot of banks which are not long dollar," says the senior fixed-income trader at the first European bank.

This funding advantage has helped US banks win over more UK clients in recent years, he says: "Because they have been long dollar through the cross-currency market they have been able to transform dollar into sterling and with sterling they have been able to generate [funding] at a level which was lower than non-US banks, thanks to the cross-currency basis and the fact that people were struggling to get the dollar."

European banks have structural advantages of their own, though. An analysis by Risk Quantum shows US dealers hold over seven times more capital against CVA exposures than European banks. This could be a result of European banks' CVA exemption for trades with EU corporates, or a sign that US banks are winning significant market share (see box: US banks' CVA lead). Either way, it should give European banks a pricing edge.

IF YOU DON'T CROSS, YOU'LL MAKE A LOSS

Cross-selling is not a new strategy in the corporates derivatives market. In order to maintain a good relationship with a large corporate name, banks have long been known to offer cheaper prices on some services such as derivatives, for instance, and then make up for that through fees on capital markets services such as mergers and acquisitions, or bond issuances.


Edouard Nguyen, Veolia

Although dealers disagree that cross-selling has become more aggressive in recent years, market participants point to two key developments.

One is that dealers have been pushing services such as wire transfers, cash centralisation and cross-border cash repatriation – collectively known as cash-pooling services – more than they did a few years ago.

"Before, it was a side business of derivatives and derivatives was really the product that was pushed. Over the past two years, I have noticed they were strongly interested in implementing cash pooling and managing cash pooling for large corporates," says Edouard Nguyen, former head of the dealing room at Paris-based corporate Veolia and now head of the corporates and treasuries practice at consulting firm Axis Alternatives.

The advantage is its retail banking-type nature, says Nguyen, which makes the business profitable and less risky. "This is a fee-based business model which is not correlated to market risk. Contracts are signed for three to four years, which provides more visibility and stability in the medium term. Also, it's less capital intensive than derivatives as banks are not lending their balance sheet," he says.

"And switching from one cash-pooling bank to another is a complicated and time-consuming process, as you need to change a lot of bank accounts, configure new standing settlement instructions, sometimes even change systems," he adds.

The head of structuring at one global bank says that recently, larger clearing banks have been laying the groundwork to combine cash sales and corporate sales teams to boost cross-selling activity.

"It used to be that cash management sales and derivatives sales were split. I hear that teams are being combined, so it's possible that these banks are considering returns from specific clients on a combined basis and the deposit or cash management side plays a role in the required returns they need for the derivatives business," he says.

Greater competition

Evolving market practices have exacerbated the race to the bottom for swap prices. For instance, the number of dealers corporates ask for quotes has increased. In a 2010 Risk.net survey of 40 corporates the average number of banks being asked to quote on a trade was two to three – none of the participants asked more than five dealers. Hedge advisers say it's now common to put 10 banks in competition for a swap.

The head of the corporates business at the second European bank says the increasing involvement of hedge advisers such as Philadelphia-based Chatham Financial has put a heavy squeeze on their margins.

"What Chatham does is go, 'So, we went to five banks. The first two had the lowest price, we think we can drive the other three into that price. Why don't we give the first two 50% of the trade and tell them they can compete with the third and fourth and fifth bank on the other 50% of the trade," he says. "They set up this dynamic that gets even the best pricers having to lower their price. It's ridiculous. I think people are caving versus holding to return limits and their return hurdles."

Brian Conly, managing director at Chatham Financial, says it understands banks need to make a fair return on capital and risk: "However, we believe that the profitability on derivatives products should be transparent and match the size of returns seen on the underlying debt/risk, not be drastically higher, as is often the case."

> "It feels like we are seeing more and more dealers coming back into this business"
>
> Senior rates structurer

Dealers say corporates are also increasingly syndicating the market risk and credit risk elements of most of their financing transactions – including cross-currency swaps – separately using a larger number of banks. This enables banks to compete aggressively on the credit risk elements of the swap's price.

"[In] some recent syndications, we have been told, 30 banks were called in the syndication process for credit. There are a large number of banks involved, including Japanese banks," says a structurer at a large global bank. "It used to be only on illiquid or large financings… it has now become a much more widespread approach… which means in our view it has decoupled the market risk component from the XVA component, potentially allowing more competition explicitly on the XVA terms."

Dealers have also become better at netting off their XVAs by viewing prices at portfolio level rather than at the trade level. For instance, two trades with the same corporate where funding costs can offset will make the second trade cheaper for the corporate. Interdealer hedges can also be chosen in a way that XVA costs offset at the netting set level.

"The target is to be more aggressive. It's a chicken-and-egg process… it's not because people have significantly improved the models that banks can be more aggressive. It's because banks tend to be more aggressive that people are thinking, 'How can we think about the models differently?" says the senior trader at the first European bank. ■

Previously published on Risk.net

US BANKS' CVA LEAD

The median credit valuation adjustment (CVA) capital charge for US global systemically important banks (G-Sibs) was 7.7 times larger than for European banks at end-2017, reflecting different implementations of Basel capital rules between the two jurisdictions.

The median capital charges for the US and European Union G-Sibs were $2.1 billion and $275 million, respectively. Aggregate CVA charges for the eight US G-Sibs were $16.3 billion, compared to just $3.7 billion for the 12 EU G-Sibs at end-2017.

Huge differences in the size of the charges can also be discerned between EU and US lenders within their respective G-Sib 'buckets', which group banks by systemic risk.

Within bucket three, for instance, Bank of America Merrill Lynch and Citigroup reported charges of $2.7 billion and $3.1 billion, respectively, whereas the two European banks in the group – Deutsche Bank and HSBC – reported charges of just $620 million and $760 million, respectively.

Nor do the differences correspond to the size of the banks' derivatives holdings. Among bucket two G-Sibs, for example, Barclays reported $13.4 trillion in over-the-counter derivatives notionals and Goldman Sachs $22.5 trillion at end-2017, meaning the US dealer's portfolio was 68% larger than the UK bank's.

However, the CVA charge for Goldman was a whopping 888% larger than for Barclays, at $3.2 billion compared with $324 million.

CVA accounts for the risk from mark-to-market losses to non-cleared derivatives due to a deterioration of counterparty credit quality. Basel Committee on Banking Supervision standards require banks to hold regulatory capital against this risk.

These standards were overhauled last year, but the changes are yet to be implemented nationally. Currently, two approaches can be used to generate the capital requirement. Under the advanced approach, banks are allowed to use a value-at-risk model to estimate losses to a 99% confidence level over a 10-day time horizon.

Two sets of results – one produced using the previous year's data and another using data corresponding to a stressed period – are summed and multiplied to produce the final charge. The standardised approach calculates the charge taking into account the external credit rating of each counterparty as well as the effective maturity and exposure-at-default of each position. Under Basel's new framework, only the standardised approach and a cruder variant will be available – internal models will be barred.

European law waives the CVA capital charge for EU banks' trades with corporations, pension funds, and sovereign entities, whereas the US law does not. The European Banking Authority has made repeated attempts to remove the exemption, which was originally championed by members of the European Parliament.

CVA charges for US banks should be higher than for their EU peers on account of their larger non-cleared derivatives portfolios, as Risk Quantum analysis relates. Non-cleared positions made up 53% of US G-Sibs' derivatives portfolios at end-2017, compared to 40% at eurozone dealers and 46% at UK banks.

However, the differences in charges are so massive that something else must be at play – namely, the effect of the EU's CVA exemptions. US bank lobbyists have long bemoaned the set of waivers allowed their European peers, and blame them for creating an unlevel playing field in non-cleared derivatives. The data appears to justify their complaints.

European authorities have themselves calculated that removing the exemptions would triple EU bank CVA capital at a stroke. Such large discrepancies may strengthen the hands of standard-setters at the Basel level when it comes to getting the EU to scrap the exemptions following the introduction of the revised CVA framework.

EU banks' CVA capital to triple if exemptions axed

Seven banks would incur 200bp-plus hit to capital if long-standing waiverswere repealed, says EBA. Writes Louie Woodall

European banks would see their credit valuation adjustment (CVA) risk capital charges more than triple if transactions currently exempted from the requirements were removed, a European Banking Authority (EBA) study shows.

The median EU bank would see its current CVA risk capital charge multiplied 3.06 times if intragroup transactions and trades with corporates, pension funds, and other non-financial entities currently spared the requirements were included in the calculations.

The survey also revealed that for 60% of banks polled, their current CVA risk charge accounted for less than 1% of their total Pillar 1 capital requirements. Just 10% of dealers said that it accounted for 4% or more.

With exempted transactions reintegrated, however, 38% of banks would have CVA charges making up 4% or more of their total requirements.

The EBA also estimated the impact on common equity Tier 1 capital of exempted transactions coming into scope of CVA capitalisation. It found that this would have a more than 200 basis point impact on seven banks' core ratios.

If half of the currently exempted CVA risk charges were thrust on banks today, the watchdog estimates the sector would be undercapitalised for this risk to the tune of €132 million. If 70% of the exempted charges materialised, the amount would be €192 million.

What is it?

The European Union's Capital Requirements Regulation (CRR) allows EU banks to waive CVA capital charges for non-cleared trades conducted with certain types of counterparties – including corporations, pension funds, and sovereign entities. The exemptions were put in place to protect non-financial counterparties from a surge in hedging costs.

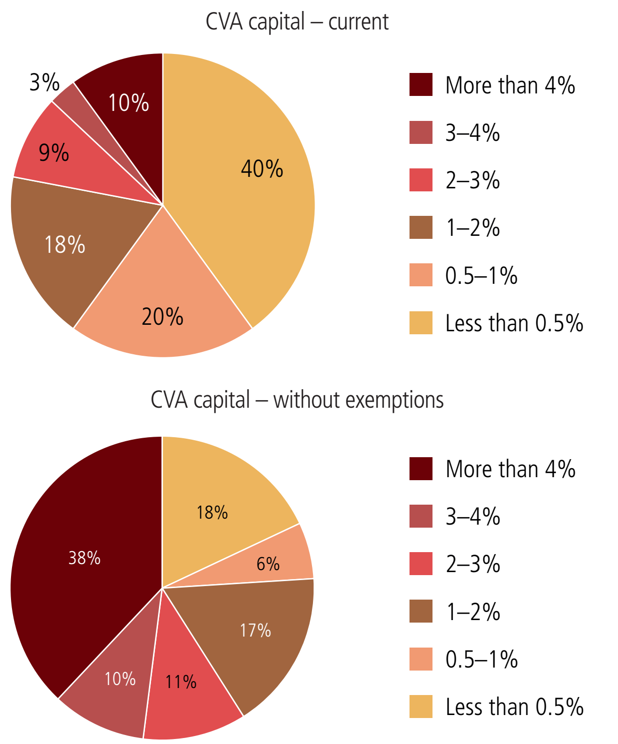
The EBA monitors the impact on bank regulatory capital of transactions exempted from the CVA risk charges. The latest results were produced from a survey of 169 EU banks, which were asked to calculate what their CVA capital charges would have been if CRR-exempted transactions had been reintegrated. Data was provided as of December 31, 2016.

Why it matters

Data-gathering exercises like the EBA's can be used as ammunition in forthcoming policy debates within the EU on whether CVA exemptions should continue to apply under revised capital requirements regulations.

The EBA itself recommended scrapping the exemptions in 2014, and issued proposals in 2015 on how the Basel standards on CVA risk could be tweaked. Yet it backed down from this position in 2017, in part because the Basel Committee nixed internal modelling for CVA capital in March 2016, leaving only the two standardised approaches on the table. Estimates suggest these approaches would impose CVA charges double those under the internal model approach.

CVA risk capital as % of total Pillar 1 capital requirements



Source: EBA 2016 CVA risk monitoring exercise

Ultimate authority for either upholding or removing the exemptions rests with the European Commission, Parliament, and Council.

Partisans on both side of the debate can use the authority's findings to bang the drum on CVA risk capital. The evidence suggests the impact of reintegrating exempted transactions would be slight on the banking sector as a whole, but that certain institutions could be crippled by drastically higher capital costs. Whether macro- or micro-prudential considerations dictate the outcome of this debate remains to be seen. ■

Previously published on Risk.net

>> Further reading on www.risk.net

- EBA shelves CVA charge plans after twin defeats *www.risk.net/4642101*
- European banks tire of CVA guessing game *www.risk.net/5129931*
- View all bank stories *www.risk.net/risk-quantum/banks*

XVA: Back to CVA?

Fundamental questions on CVA remain unanswered, says mathematical finance head Damiano Brigo

While many professionals are chasing the latest derivatives valuation adjustments (XVAs) and trying to introduce new ones, we should make sure we deal with the fundamental ones first.

Many basic issues with credit valuation adjustment (CVA) – the first of the family – still remain unaddressed. We should be cautious in considering it dead due to the effects of margining for non-cleared swaps, especially given the blunt methodology adopted for the industry's standard initial margin model (Simm), and we should also pay attention to potential double counting in XVAs due to non-linearities.

CVA answers the following question: "How much discount do I get on the price of this deal due to the fact that you, my counterparty, can default?" It has traditionally been calculated with a risk-neutral valuation approach (Brigo and Masetti, 2006), with all the pros and cons discussed in my previous column (*Risk* Magazine, February 2018).

But despite having been around for years, a question remains: can CVA really be hedged? Challenges range from finding good liquid instruments for default probabilities and recoveries, to modelling the dynamics and option prices of complex netting sets with option maturities given by the counterparty random default time. Wrong-way risk is also quite model-dependent and hard to assess (Brigo and Pallavicini, 2008).

Then there is debit valuation adjustment (DVA), which is the CVA seen from the other counterparty's perspective (Brigo and Capponi, 2008). DVA answers the question: "How much markup do I need to pay over the price of this deal to my counterparty due to the fact I can default?"

The mark-to-market value of DVA goes up when a company's credit quality goes down. Companies could therefore profit from their debt deterioration – we have seen banks report $2.5 billion DVA gains in a quarter in the past.

Given that a party cannot sell protection on itself, DVA is notoriously difficult to hedge. It is typically hedged by proxy, which is not ideal when jump-to-default risk is included in the picture. Furthermore, bilateral CVA and DVA introduce a first-to-default time that embeds a default correlation that is difficult to hedge. The Basel Committee on Banking Supervision hasn't recognised DVA, whereas the accounting standards do.

Some ask will margin and clearing kill CVA and DVA? In a 2014 paper on bilateral counterparty risk valuation (Brigo, Capponi and Pallavicini, 2014), it is shown that even under continued collateralisation, contagion and gap risk at default may lead to a residual CVA that in some cases is as large as CVA as for the uncollateralised trade. In a 2014 paper on cleared and bilateral swap pricing (Brigo and Pallavicini, 2014), we also study liquidation delays coming from possible disputes.

The additional initial margin might limit the problem, but the blunt methodology adopted by the Simm implies it does not always address the real gap risk, which will be always quite model-dependent (Brigo and Pallavicini, 2014).

FVA is the next in the XVA family. This accounts for all the borrowing costs and lending benefits one faces in servicing the trade accounts. It can be sizeable – JP Morgan, for example, declared an FVA of $1.5 billion in a single quarter in 2013.

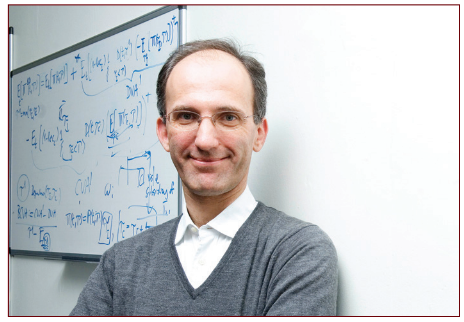
FVA is linked to the CVA and DVA that come from the external borrowing and lending operation of the bank, so in a way it is also driven by credit risk. These three XVAs may introduce non-linear features in valuation when borrowing and lending rates are not equal, or when replacement closeout at default is used in case of early default.

The mathematical tools needed in this case range from semi-linear partial differential equations to backward stochastic differential equations. These tools are quite advanced and rarely used in the industry. They can also be used to assess the cost of trading via a clearing house or a standardised credit support annex with variation and initial margin (Brigo and Pallavicini, 2014). Approximating non-linearities by linearising can lead to double counting, and the issue needs to be investigated further.

The final valuation adjustment is capital valuation adjustment (KVA). Initially, a classic replication approach was proposed in 2014 (Green, Kenyon and Dennis, 2014). However, as I mentioned in the February column, there is a different approach: in a 2017 paper (Brigo, Francischello, Pallavicini, 2017), we propose an indifference approach for KVA, finally moving beyond the continued and unrealistic stretch of risk-neutral valuation and replication.

However, while further research on KVA and XVA is needed, we should not forget the unsolved, fundamental challenges around CVA. ■

Previously published on Risk.net



Damiano Brigo is head of the mathematical finance research group at Imperial College, London

REFERENCES

Brigo D and Masetti M, 2006
Risk Neutral Pricing of Counterparty Risk
Counterparty Credit Risk Modelling: Risk Management, Pricing and Regulation (edited by Michael Pykhtin)
Risk Books

Brigo D and Pallavicini A, 2008
Counterparty Risk and Contingent CDS under correlation
Risk February, 2008

Brigo D and Capponi A, 2008
Bilateral counterparty risk with stochastic dynamical models
arXiv.org

Brigo D, Capponi A and Pallavicini A, 2014
Arbitrage-free bilateral counterparty risk valuation under collateralization and application to Credit Default Swaps
Mathematical Finance, Vol 24, No 1, pages 125–146

Brigo D and Pallavicini A, 2014
Nonlinear consistent valuation of CCP cleared or CSA bilateral trades with initial margins under credit, funding and wrong-way risks
Journal of Financial Engineering, 1 (1):1–60

Green A, Kenyon C and Dennis CR, 2014
KVA: Capital Valuation Adjustment by Replication
Risk 27(12)

Brigo D, Francischello M and Pallavicini A, 2017
An indifference approach to the cost of capital constraints: KVA and beyond
arXiv.org

XVA management

Challenges and solutions

Amid a lack of established best practice on how to manage and calculate XVA, for many firms standardision is priority. In our XVA management forum, a panel of industry leaders discusses key topics, including the effect a changing regulatory landscape is having on XVA management, the potential impact of cloud computing and web-based technology, and industry-wide limitations with XVA calculation and how the panellists' respective organisations are addressing them

Stéphane Rio, Founder
and Chief Executive, ICA
www.the-ica.com

What are the biggest concerns currently about how the industry manages XVA?

Stéphane Rio, ICA: Beyond the lack of standardisation and transparency, which are key issues for the industry, each bank has also its own challenges. Typically, there are two areas that are often contradictory, but always intertwined, in which the current situation seems unsatisfactory for XVA desks: speed – or efficiency – and acute understanding.

A typical illustration of concerns with speed is traders having to assess – in as close to real time as possible, which more often ends up being done in minutes rather than in a couple of seconds – the impact of a transaction on XVA. Add in the resulting changes to hedges and the entire process often becomes a daunting task.

In terms of understanding, whether profit and loss (P&L) explanations or sales attributions, determining the source of a change is similarly difficult.

Often, to resolve the first problem, solutions have compromised the second, and vice versa.

How difficult is it to make centralised XVA decisions with siloed trading desks?

Stéphane Rio: XVA desks have the unique feature of being true cross-asset desks. In particular they are, by nature, forced to rely on data – trades, counterparty and market data – coming from all other desks in the bank. This raises questions around the heterogeneity and quality of this data, which will drive important decisions by the XVA desk. Appropriate processes and controls have to be put in place to mitigate this risk.

A second aspect is organisation. When it comes to pricing a client trade, there are several parties involved in building the final price, including the trading desk (risk-free price), the XVA desk (XVA margin) and the sales desk (sales margin). A robust XVA system must account for an efficient sales-pricing workflow, including for all actors' interactions, and ensure a timely response to the client.

How significant a concern is needing to make hard and fast decisions on XVA in a changing regulatory landscape?

Stéphane Rio: Such a landscape often translates into new and potentially complex system requirements. However, in practice, system evolutions are very slow, because of system landscapes in banks that are still quite monolithic. In the case of traditional system vendors, upgrading to attain new functionalities is often a large project that can take months, if not years.

For banks undertaking internal development, there is a need to separate tasks into independent modules to gain agility and adapt more quickly to the required evolution. For banks using vendors, it is time to think of adopting Software as a Service (SaaS) – adjusting to regulatory changes will only be a matter of testing the new results or connecting the results to the internal workflow, and can be achieved in just a few weeks at minimal cost.

What are the current limitations for the industry when calculating XVA?

Stéphane Rio: Big compute and big data. Banks inevitably bump into large parallel compute issues and the capacity to manipulate a very large amount of data.

Often these hurdles are resolved through approximations – thereby avoiding a full revaluation – and through discarding intermediate results, which creates additional problems:

- Model validation challenges – in particular when models from the risk department differ from the front-office models.
- The inability to save intermediate results complicates the analysis of results and requires full recomputing for each incremental pricing.

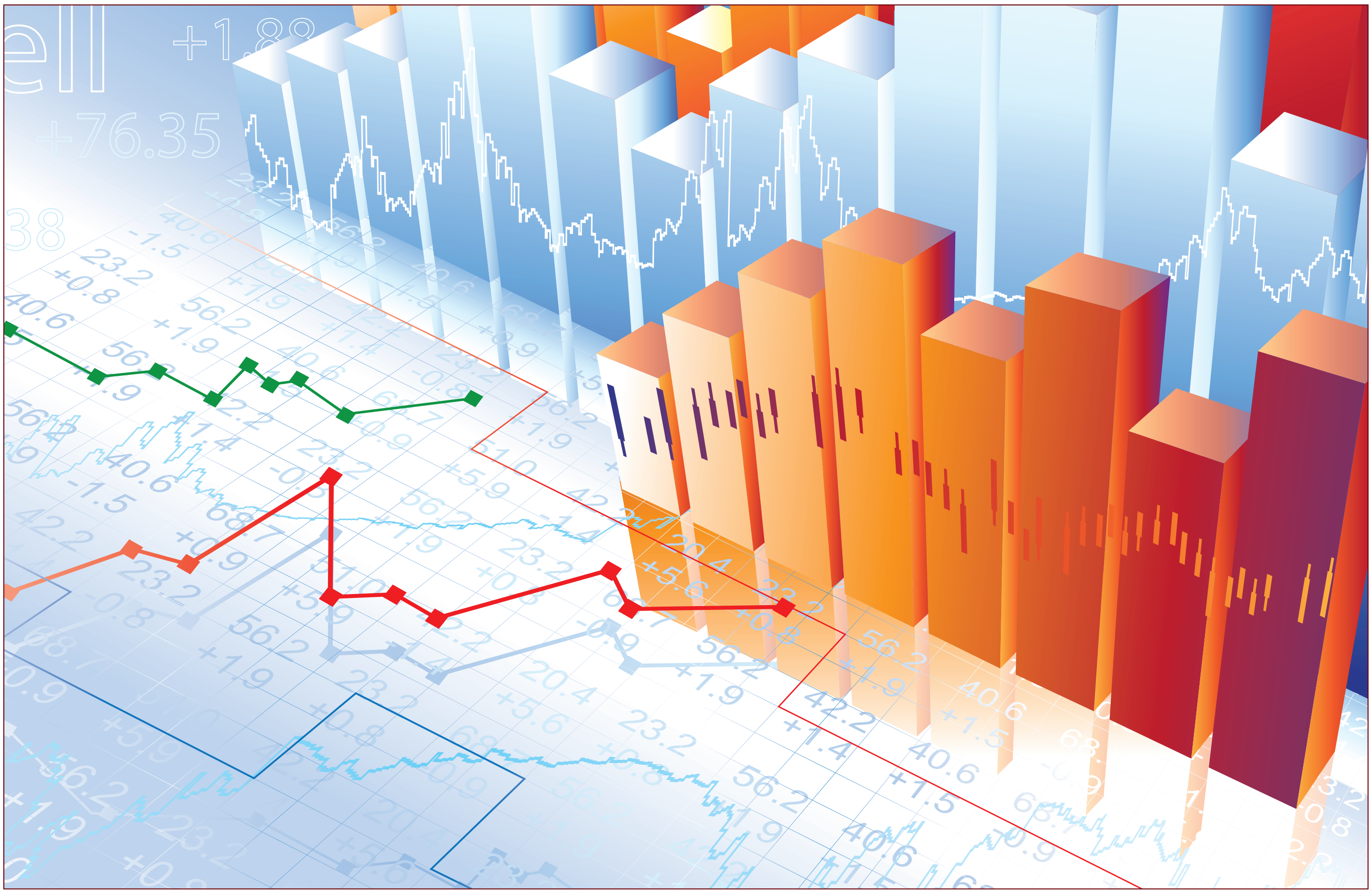
Furthermore, access to sufficient compute power is critical. However, when using a finite quantity of in-house servers, XVA desks lack compute power for night batches – for in-depth sensitivities, cross-gamma or stress-test analysis, for example – and must bear the cost of 'sleeping resources' for most of the day.

In what way would your organisation address these limitations?

Stéphane Rio: Those limitations are what ICA is focusing on. We don't believe models are the issue, but rather the implementation of models in the right architecture and infrastructure. In delivering a fully serviced solution, we deal with the full processing chain and have made it our speciality to tackle big compute issues on behalf of banks.

ICA has embraced all the recent big data and cloud technologies to address those issues and benefit from the following:

- Combining the business, digital technologies and quant expertise into a single team



- Leveraging the cloud for all non-confidential calculations, allowing the majority of compute to be fully scalable and elastic
- Being able to save and manipulate enormous quantities of data in a database, rather than in memory.

This allows us to:

- Optimise the distribution of calculations (avoid redundant calculations, minimise input/output, and so on)
- Compute cross-gamma or stress-test scenarios on demand. More generally, compute what you need rather than what you can or what you did the day before
- Access all intermediate calculations in real time, facilitating result investigation and generation of what-if scenarios in real time – pre-trade pricing, post-trade optimisations, changes of credit support annex terms, central counterparty upload, and so on.

How influential could cloud computing and web-based technology be in transforming the calculation of XVA and the management of XVA data?

Stéphane Rio: Using the elasticity and scalability of the cloud to run massive computations is now largely recognised as a required feature of new-generation XVA solutions (see *When a cloud can light the way*, page 18). The challenge, however, lies in how secure this data and these processes will be. Regulators have issued sensible guidelines, and additional security can be added to that. For instance, ICA's process will strip any confidential information – such as notionals, counterparties or netting sets – out of deal descriptions before anything is sent to a public cloud.

Web-based technologies are also key contributors to the reshaping of the landscape for XVA and derivatives pricing and risk calculations. This is a core aspect of ICA's value proposition; we support banks in implementing those technologies to modernise their architectures. Interestingly, digital innovators have also presented a new way to approach processes and build setups through the flexibility and efficiency of SaaS, which is surely the future of financial software.

Could a series of deregulations render investment in XVA calculations pointless?

Stéphane Rio: I don't think one should take bets on whether there will be more or less regulation, as it is ever-changing. But whether because of changing regulation or deregulation, banks should be agile in their development processes and vendors' upgrades.

Which XVAs are causing the most headaches at the moment, and why?

Stéphane Rio: A typical example of XVA metrics that is highly challenging – or will become so in the near future – is the CVA capital charge under the Fundamental Review of the Trading Book's standardised approach. Solely for the spot capital, it requires the calculation of the CVA sensitivities across pretty much all risk factors – rates, volatility, credit, and so on.

For full KVA pricing, this process must potentially be repeated for every projected time step. Lastly, trade allocation for that metric will involve even more complex calculations – multidimensional and non-linear solving – so banks must calibrate compute power accordingly and consider strategies and clarity about allocations of capital by trade, desk and business line. ■

When a cloud can light the way

Proponents of cloud computing say it is the only way to fully meet the demands of FRTB, XVA and other changes. By Nazneen Sherif

Need to know

- Financial institutions are looking to cloud computing to carry out their complex calculations, marking a shift from investing heavily in in-house systems that are expensive to develop, maintain and upgrade.
- The search for additional computational capacity stems from regulatory demands such as FRTB and derivatives revaluation efforts, among others.
- Because firms can access much more hardware on the cloud, speedup of anywhere between 10 and 1,000 times is possible.
- Proponents argue cloud services are more economical, and can lead to savings on IT budgets.
- Others report having to spend more on security to protect sensitive data hosted on the cloud.

Technology is a fickle thing. Older, less-efficient technologies that once revolutionised an industry are constantly being replaced by newer, better-performing solutions. For risk managers weighed down by the computational burden of new market risk rules introduced by the Fundamental Review of the Trading Book, these technological advances cannot come soon enough.

That is why many large dealers are exploring cloud computing as a way of overcoming the perceived limitations of in-house systems. Those that have tried and tested cloud computing report calculations can be sped up by factors of tens, if not thousands, and IT costs slashed.

But exporting vast quantities of sensitive data to a third party comes with its own risks, and the necessary expense incurred in protecting that data threatens to chip away at the much-vaunted cost savings.

"One of the biggest computation-intensive use cases we have in the industry is the generation of full revaluation-based P&L vectors for the FRTB internal models approach," says Suman Datta, a director in the portfolio quantitative research team at Lloyds Banking Group in London. "The number of scenarios and calculations required are enormous and the cloud gives firms the flexibility and scalability to deliver this."

John Kain, a business development manager in the capital markets team at AWS, Amazon's cloud services provider, expects regulatory reporting under FRTB to lead to a tenfold increase in portfolio risk calculations.

FRTB is not the only recent development to have stretched bank IT resources: derivatives valuation adjustments, or XVAs, have proved a comparable drain on dealers' computational capacity. Portfolios containing thousands of trades require thousands of revaluations to be carried out at each time point in the future.

"I have heard a number of dealers saying they have had instructions not to increase their internal power or capacity any further because there are more and more requirements from XVAs and the FRTB," says Stephane Rio, who runs The Independent Calculation Agent, an XVA and FRTB pricing firm that uses cloud technology. "They are turning to the only long term viable strategy for extending their capacity to meet those needs, which is leveraging the scalability and elasticity of the cloud."

Three European dealers confirmed to *Risk.net* they will be transitioning to the cloud within the next year to manage their FRTB and XVA calculations. A spokesperson for Nordea said the bank is currently in late development and early test phase to use the technology for end-of-day and XVA calculations.

Cloud computing is a way of sharing hardware and software via a network such as the internet. The information is stored on physical servers maintained by a cloud computing provider. The main service providers in the financial services industry currently include AWS, the Google Cloud Platform and Microsoft Azure.

The advantage of sharing resources is that firms accessing the cloud need not maintain their own hardware and data centres on a permanent basis. Instead, they can access the platform from any location only when they need it, based on a pay-as-you-go structure. The scale of the infrastructure offered by these large vendors also means dealers can access thousands of cores – both traditional central processing units (CPUs) and more advanced graphics processing units (GPUs) – on an *ad hoc* basis.

This on-demand scalability is a principal driver for cloud computing adoption among financial institutions.

"In the past, you would have to make the long-term decision, saying 'I expect the business to grow X amount every year and hence this is the number of cores I need to buy in advance'. Whereas with the cloud, theoretically, you have that option where if you need 1,000 more cores tomorrow, you have the ability to do that easily," says a risk manager at a European bank.

There are two ways of accessing cloud services. One is by sharing all resources with other firms signed up to the service, which means the number of cores available to a firm is not guaranteed. The second – and probably more preferred – option is using a 'hybrid' model where an institution reserves a fixed amount of cores so they can ensure they do not run out of machines at any given point in time.

For the European bank, which is currently in the development phase, the reason for moving to the cloud was being able to do its risk and pricing calculations more frequently – especially for more granular risk management.

> "It doesn't make sense to own additional data centres. It is impossible to maintain. The cloud is the only option"
>
> Risk manager at a European bank

"Our market risk front-office departments would like to have more risk measures available in real time done at a much higher frequency, rather than just once a day or three times a day; they now want six times a day – so they want to be on top of the risk they are running," says the risk manager.

The bank confirmed it will be moving its XVA pricing, risk management, stress testing and prudential valuation calculations to the cloud within the next year.

Another large US bank is already using cloud computing to run quarterly capital and stress runs, but confirmed it is exploring further use cases.

Buy-side firms are also using the technology to improve computationally intensive calculations such as portfolio optimisation, where returns are optimised given the risk or volatility of the assets. This process includes estimation of the correlations between the assets to factor in how they move together. This becomes extremely cumbersome for portfolios with a large number of assets. For instance, for a portfolio of 100 assets, one would need to estimate a 100 by 100 correlation matrix.

In addition, correlations only capture the linear relationship between how assets move. Modelling non-linear relationships may be more accurate, but can be very intensive to run computationally.

"People could just do mean-variance optimisation but now you can optimise in various ways, given the computational power. Before, people used to use correlations to measure their covariance matrices, now you can use other types of measures of non-linear relationships," says one hedge fund manager in New York. "When you cross the line between linear relationships and non-linear relationships, that's when computational needs accelerate at an exponential basis."

Another application the firm is considering running on the cloud is machine learning.

Both banks and asset managers have been using machine learning techniques to explore a number of applications in finance including trade execution and model validation. Because machine learning techniques work by running millions of simulations to identify patterns in data and choosing the best course of action, high computing power is necessary to be able to employ them.

"Computational techniques like machine learning, which requires a lot more CPU or GPU usage, is probably one of the biggest applications [of cloud computing]," says the hedge fund manager.

One hand giveth…

Faster calculating speeds, coupled with the ability to scale up capacity at short notice, can result in dramatic cost savings, users say. Typically, a large dealer would require about 10,000 to 30,000 cores for five to seven hours a night to run their complex pricing and risk calculations, says a risk manager at a second European bank. Each CPU has about eight cores and costs around £1,000. GPUs, on the other hand, have about 3,000 cores each and could cost between £2,000 and £8,000. GPUs also entail additional maintenance and development expenses, not to mention power consumption, space and cooling requirements. For firms that have invested in their own technology, these machines sit idle for about half the time, says one large dealer.

While some banks have been able to mitigate that expense by switching to techniques such as adjoint algorithmic differentiation, which can be used as a relatively cost-free alternative to GPUs to calculate risk sensitivities for simple products, a significant portion of pricing and regulatory calculations still need raw computing power.

"If you have extremely complex computations that require lots of GPU parallelisation, we can now actually do that without spending tonnes of money setting up actual boxes in a data centre and paying for renting shelf and data centres – it's really revolutionary across the entire IT industry," says the New York-based hedge fund manager.

The risk manager at the first European bank agrees: "It's not easy to buy a data centre. If you

WHO IS ON THE CLOUD?

Risk.net spoke to two mainstream providers of cloud computing services to financial institutions: Amazon's AWS and Microsoft Azure.

A number of dealers, insurers and stock exchanges are using the platforms to meet complex business needs.

Microsoft Azure signed up UBS to its service in 2017 to help the Swiss bank carry out its risk management calculations. Insurer MetLife also uses the service for its MetLife Integrated Actuarial Modeling Environment, running complex simulation models. The saving for MetLife is estimated to be between 45% and 55% in infrastructure costs.

Users of Amazon's cloud service include asset manager Talanax, Spanish insurer Mapfre, Starling Bank, Liberty Mutual Insurance, Thomson Reuters, DBS Bank and Finra.

Financial services provider Aon Securities uses AWS to run financial simulations to value and manage insurance retirement products using GPUs.

"Aon has been able to lower the calculation and total reporting process time from 10 days to 10 minutes," says John Kain at AWS.

Nasdaq, Robinhood, London Stock Exchange and Aire also use the AWS service for analytics applications to optimise for cost effectiveness, scalability, and reduce processing times for risk-related workloads. The risk applications include actuarial calculations, CCAR, FRTB and Solvency II.

want to scale up, you need to plan a year in advance before that transaction gets done, whereas with the cloud you should be able to be much more agile in your ability to do that."

"It doesn't make sense to own additional data centres. It is impossible to maintain," says the risk manager at a second European bank. "The cloud is the only option."

Renting capacity on the cloud costs around £3.50 for 24 GPUs per hour. Cloud computing providers also update their hardware regularly, which means banks would not face the additional expense of updating their technology.

UBS, which is using the Azure service, says it was able to bring down its running costs by 40% after moving its risk management platform to the cloud.

Not everyone is convinced there is an overall cost saving, however. Cheaper processing capability does not necessarily equate to lower IT spend – especially when the amount of required calculations is increasing.

"The primary goal is not to reduce the compute budget but rather to make computing more flexible and economical. In fact, cloud computing might lead to an increase in our compute budget by making more calculations feasible and economically worthwhile," says the Nordea spokesperson.

…the other taketh away

One reason banks have been slow to use cloud service providers is because of concerns around the sharing of confidential client and portfolio information on an external platform. To carry out pricing and risk management calculations on the cloud, dealers must share sensitive data, which could include client information for portfolio-level calculations such as XVAs. Some argue the cost of securing this data on the cloud neutralises the gains from the pay-as-you-go arrangement.

"The drawback or challenge with using outside cloud solutions for banks, and XVA in particular, is around securing data, and especially sensitive client information. The setup around this is, and should be, really strict," says a CVA trader at a third European bank.

A risk manager at an Australian bank agrees: "What we save in terms of grid costs we gain in terms of additional security costs. For example, our IT and security teams at all times are making sure the information we are posting to the cloud is appropriately protected – it's not a free pass."

Wary of relying on a service provider's level of protection, some banks are considering withholding certain types of data from the cloud, to help plug any security gaps.

"We try not to provide information on four elements: client identity, our market data, our positions and also our own internal modelling," says one XVA quant based in Europe.

This, however, is easier for pricing of individual trades. When it comes to more complex calculations such as XVAs, portfolio information is needed to aggregate results, which cannot be done without client data.

One way to get around this is to alter the data being sent to the cloud.

"Pre-processing can be done to hide information. For example, if you have a swap of a given notional you are not obliged to send the exact notional amount, you can divide it by a certain factor and rescale it back," says the XVA quant.

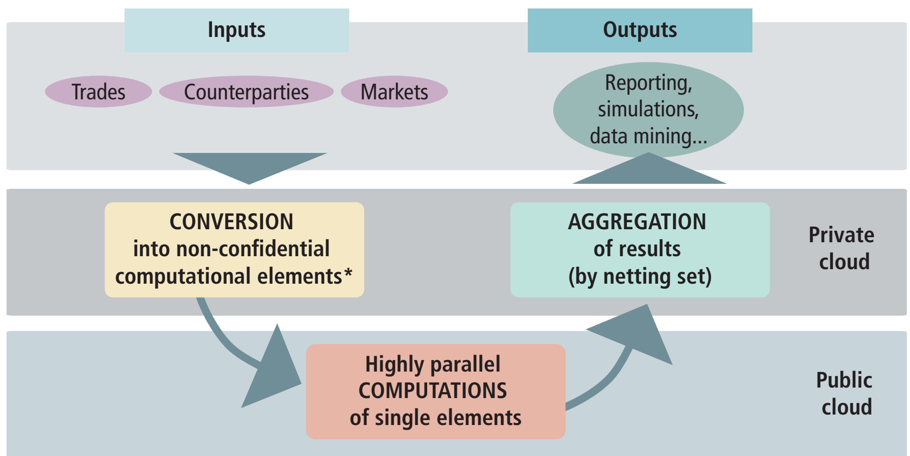
Another solution is to split up computations into two steps, so the part that is run on the cloud does not contain any confidential information. Independent Calculation Agent conducts 85–90% of a portfolio computation on the cloud, but the remainder takes place either within the bank or ICA's own data centre in a private and secured environment (see figure 1). The results are aggregated in a bank's internal system.

"The only thing you will do locally are very simple operations, but very critical in terms of confidentiality," says Rio at ICA.

Alexander Sokol, chief executive officer of vendor CompatibL, downplays the security concerns of mainstream cloud providers such as AWS, Azure and Google.

"Because of multiple security protocols such as encryption of data at rest and encryption of data in transit, a well-managed public cloud from a mainstream provider such as Amazon is more secure, or as secure as private data centres," he says.

1 Splitting computations on the cloud to protect sensitive data



*Trades are transformed into independent computation elements such as normalised cashflows and instruments

Source: The Independent Calculation Agent

When disaster strikes

Security and cost considerations aside, proponents of cloud computing argue the technology offers benefits in terms of disaster recovery and data replication.

"For a fund like ours, it is a much safer way for disaster recovery. We have multiple virtual private clouds in different geographical regions, so we do half of our computations in northern Virginia and half in Ohio," says the New York-based hedge fund manager. "So even if Northern Virginia were to suffer a failure, it's literally seamless for us to switch to Ohio fully, rig up copy instances and rig up new machines. We can do all that in 15 to 20 minutes as long as we don't lose internet access."

But what happens if the disaster afflicts the cloud services provider? Overreliance on cloud platforms could mean when these firms experience a shutdown, dealers might find themselves short of resources.

As the risk manager at the first European bank says: "You are taking a risk in that you are dependent on an external provider, and what happens if that external provider goes down because of financial troubles? Then you have tied your fortunes to those cloud providers." ■

Previously published on Risk.net

>> Further reading on www.risk.net

- Quants turn to machine learning to model market impact *www.risk.net/4644191*
- Model risk managers eye benefits of machine learning *www.risk.net/4646956*
- AAD vs GPUs: banks turn to maths trick as chips lose appeal *www.risk.net/2389945*

Machine learning

Not just for the buy side

Sell-side quants develop machine learning technique to optimise margin costs. By Nazneen Sherif

The most common application being researched for machine learning is optimal execution. When large trades are executed in the market, it could potentially push prices in an unfavourable direction, so it makes sense that traders are keen on optimising this cost.

So far, most of the interest in applying machine learning technology to reduce trading costs has been from the buy side.[1]

However, recent research by quants from Standard Chartered shows this may be about to change.

In this month's first technical, *Evolutionary algos for optimising MVA*[2], Alexei Kondratyev, a managing director at Standard Chartered in London, and George Giorgidze a senior quantitative developer in the strats team within the same bank, propose machine learning techniques to optimise initial margin costs through trade selection.

Since September 2016, an increasing number of large dealers have been required to post initial margin[3] on new non-cleared trades with other in-scope counterparties. The initial margin has to be funded, which creates material costs as more and more counterparties become involved. This cost is typically priced into trades in the form of a margin valuation adjustment (MVA).

Market participants have estimated initial margin funding requirements under the regime to be close to $1 trillion. As a result, a number of margin compression and risk optimisation solutions have popped up to reduce margin funding costs, each promising more significant margin reductions than the last.[4]

In their paper, Kondratyev and Giorgidze, apply two machine learning algorithms – a genetic algorithm and a particle swarm optimisation (PSO) – to reduce margin costs over the life of the portfolio, while keeping the market risk exposure of the portfolio the same.

This is not easy. Because there are many parameters that evolve over time, the problem becomes heavily non-linear – that is, traditional optimisation methods do not work in simulating and reducing MVA.

This is where the benefits of machine learning come in, says Kondratyev.

Both algorithms used by the quants belong to the class of so-called evolutionary algorithms that run multiple iterations of chromosomes, tweaking one or two genes at a time, to find the most beneficial mutation. The StanChart quants apply the same principle to MVA.

Here the objective function, which defines the quantity to be maximised or minimised, is incremental MVA on a portfolio of trades. The chromosomes represent individual trades and the genes are trade details such as direction, notional size and currency.

"One could take a snapshot, look at portfolios, as of say, end of day, and then one would see that if a couple of trades are added into the portfolios, the Simm initial margin profiles would be flattened out and MVA would be reduced. Then, a proposal can be made to the trading function and it's up to them to decide whether to go ahead and execute these trades," says Kondratyev.

The techniques have been in use in Standard Chartered since the beginning of the year. The quants did not reveal what their MVA savings were, but they said it would be significant enough to justify the cost of development of the technique,

> "From generation to generation, our calculation becomes better and better, until finally we evolve to a population of solutions that converges to a global minimum of our objective function"
>
> Alexei Kondratyev, managing director, Standard Chartered

also factoring in future increases in MVA as more counterparties come in scope of the margin rules.

"The algorithm works by finding such values in our solution, which is a trade or pair of offsetting trades that would minimise our objective function, and our objective function is incremental MVA, so we try to find the most negative incremental MVA," says Kondratyev.

The algorithm randomly generates trades by tweaking one or two trade details each time, and keeps those that reduce MVA and discards those that increase MVA – similar to how evolution works. The genetic algorithm is used when trade details are discrete or categorical, such as currency for example. In this case all variables need to be discretised – for instance tenor, can be discretised by month. PSO is used when they are continuous – notional size, for instance, is a continuous variable.

"From generation to generation, our calculation becomes better and better, until finally we evolve to a population of solutions that converges to a global minimum of our objective function," adds Kondratyev.

The resulting set of optimal trades could then be used to guide traders on what sort of trades and counterparties would help reduce their margin costs under the industry's standard initial margin model.

In the race to survive the rising costs of the derivatives business, it is not surprising that quants have started to use evolutionary algorithms to reduce what is likely to be one of their biggest costs – the funding initial margin.

Derivatives valuation adjustments, in general, are complex to model. So optimising them requires more innovative techniques.

While buy-side quants have done most of the early exploring of machine learning, Kondratyev and Giorgidze's paper is a step in the right direction towards encouraging more sell-siders to explore the technique as well – which may become crucial as MVA becomes more substantial over time. ■

Previously published on Risk.net

[1] *Day S, Quants turn to machine learning to model market impact, April 2018, www.risk.net/4644191*
[2] *Kondratyev A & Giorgidze G, Evolutionary algos for optimising MVA, December 2017, www.risk.net/5374321*
[3] *Risk.net, CFTC relief caps day of swaps margin mayhem, September 2016, www.risk.net/2469566*
[4] *Woodall L, Non-cleared swaps compression battle heats up, March 2016, www.risk.net/4008711*

Automatic backward differentiation for American Monte Carlo

Christian Fries derives a modified backward automatic differentiation (also known as adjoint algorithmic differentiation) for algorithms containing conditional expectation operators or indicator functions. Bermudan option and XVA valuation are prototypical applications. Featuring a clean-and-simple implementation, this method improves accuracy and performance. It also enables accurate 'per-operator' differentiation of the indicator function (exercise boundary)

The Monte Carlo valuation of Bermudan-like products or any valuation requiring the calculation of conditional expectations of future values (credit valuation adjustment (CVA) and margin valuation adjustment (MVA) are common examples) is a non-trivial problem. The standard solution is to use regression methods to estimate the conditional expectation, often referred to as 'American Monte Carlo'.

Another numerically intensive problem is the calculation of sensitivities (Greeks), ie, partial derivatives with respect to model parameters. This problem can be solved efficiently by adjoint automatic differentiation (see Giles & Glasserman 2006).

If the valuation algorithm involves a conditional expectation operator, the calculation of sensitivities then requires the differentiation of the conditional expectation estimator. In some cases, the differentiation of the conditional expectation may be omitted, eg, if the conditional expectation is the input of an optimal exercise criterion and the sensitivity is a first-order sensitivity (see Piterbarg 2004). However, in general, the differentiation cannot be omitted. See the numerical results below for examples.

Since the exercise criterion is essentially an indicator function, the differentiation of the indicator function is another numerically demanding problem.

We consider the application of adjoint automatic differentiation to calculate the sensitivities of a general product valuation involving a conditional expectation operator and indicator functions. The automatic differentiation of valuations involving conditional expectation has been discussed in, for example, Andreasen (2014), Antonov *et al* (2018), Antonov (2017) and Capriotti *et al* (2017).

To understand the different approaches, it is important to understand that automatic differentiation comes in essentially two different flavours: it can operate in forward mode (forward AD; sometimes called simply AD), where the derivatives are propagated forwards from the inputs (parameters) to the results (values), alongside the valuation; or it can operate in backward mode (backward AD; sometimes called adjoint AD or AAD), where one propagates derivatives backwards from the results to the inputs. The numerical performance of the forward mode scales with the number of parameters and is roughly independent of the number of results, while the performance of the backward mode scales with the number of results and is roughly independent of the number of parameters.

Adjoint algorithmic (backward mode) differentiation is the method of choice when sensitivities have to be calculated for a single result (or a few results) depending on many parameters. A striking example is the valuation of an MVA from Isda-Simm initial margins (see Fries 2019).[1]

[1] *Isda-Simm is the International Swaps and Derivatives Association's standard initial margin model.*

Applying an automatic differentiation, and an adjoint automatic differentiation in particular, to a valuation algorithm containing a conditional expectation reveals some issues. While in a Monte Carlo simulation most operators are pathwise, and differentiation can be applied on a path-by-path basis, the conditional expectation operator is non-pathwise, aggregating information from adjoint future paths. A brute-force application of (A)AD to the conditional expectation regression will differentiate the regression basis functions (see Andreasen 2014; Capriotti *et al* 2017).

However, differentiating the regression basis function can always be avoided, as long as the filtration does not depend on the model parameters.[2] Then, the differentiation of a conditional expectation is the conditional expectation of the differentiation:

$$\frac{\mathrm{d}}{\mathrm{d}x} \boldsymbol{E}(Z \mid \mathcal{F}_t) = \boldsymbol{E}\left(\frac{\mathrm{d}}{\mathrm{d}x} Z \;\middle|\; \mathcal{F}_t\right) \tag{1}$$

This result offers another striking optimisation: we may just check if $(\mathrm{d}/\mathrm{d}x)Z$ is an $\mathcal{F}_t$-measurable random variable. If that is the case, we can omit the outer conditional expectation operator on the right-hand side. Our implementation at http://bit.ly/2tBxDQK contains an automatic tracking of the measurability of random variables, which enables us to drop the conditional expectation whenever possible.[3]

If the conditional expectation operator cannot be omitted (see below for examples), the application of (1) still imposes a difficulty for the implementation of adjoint automatic differentiation. Since the AAD operates backwards through the operators, we cannot apply (1) during the backward propagation, as we have to calculate the inner derivative $(\mathrm{d}/\mathrm{d}x)Z$ first, which is only available after the backward sweep has been completed.

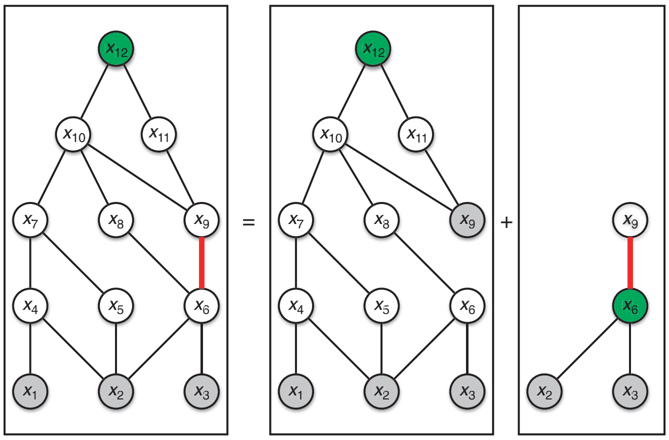
This issue may be solved by splitting the algorithm into two independent backward-differentiation algorithms, which must then be combined (see figure 1).

In Antonov *et al* (2018), the authors present a method called backward differentiation (BD), which calculates the derivatives of the arguments of conditional expectations (the so-called continuation values) alongside the valuation. This algorithm has the advantage that the derivatives can be calculated in a single sweep. To do so, they directly modify the valuation code. The name 'backward differentiation' is somewhat misleading: the BD algorithm in Antonov *et al* (2018) is propagating derivatives forwards

[2] *This is the case in supposedly all practical Monte Carlo applications: the filtration is represented by the generated random numbers.*

[3] *In Antonov et al (2018), the authors illustrate that for their algorithm the differentiation of the conditional expectation can be avoided, given the absence of path-dependency; however, this condition may be hard to check.*

1 Operator tree in which x_9 is a conditional expectation of x_6



The differentiation algorithm can be decomposed into the differentiation of x_{12}, considering x_9 as an independent variable, and the differentiation of x_9

(namely alongside the valuation algorithm). It is just that the algorithm itself goes backwards over the time steps, since this is the way a Bermudan callable is valued. While this allows us to calculate the derivative in one sweep, it means the performance and complexity of the algorithms scales with the number of parameters (see N_θ in Antonov *et al* (2018, section 7.4)), as is typical for a forward mode AD.

In contrast, we will show how to efficiently apply a true backward mode differentiation (ie, AAD backward differentiation) to a Bermudan callable, without modification of the original valuation algorithm. The trick that enables us to treat the conditional expectation in a single backwards sweep relies on the following observations:

- The application of an automatic differentiation to an algorithm containing a conditional expectation results in a linear combination of conditional expectation operators (due to (1)).
- The valuation is an expectation, that is, the last operator is an expectation operator.[4]
- For any two random variables A, B, we have that:

$$\boldsymbol{E}(A\boldsymbol{E}(B \mid \mathcal{F}_t)) = \boldsymbol{E}(\boldsymbol{E}(A \mid \mathcal{F}_t)B) \tag{2}$$

Equation (2) allows us to apply the conditional expectation to the adjoint differential. The adjoint differential is available during the backward differentiation once the algorithm reaches the conditional expectation operator.

It remains to show (2) and (1) can be applied in an algorithm featuring multiple, possibly iterative, conditional expectations, as is common in a Bermudan backward algorithm.

Setup and notation. Given a filtered probability space $(\Omega, \mathbb{Q}, \{\mathcal{F}_t\})$, we consider a Monte Carlo simulation of a (time-discretised) stochastic process, ie, we simulate a sample path ω_i of sequences of random variables. Assuming the drawings are uniform with respect to the measure $\mathbb{Q}$, this allows us to approximate the unconditional expectation $\boldsymbol{E}^{\mathbb{Q}}(X)$ of a random variable X via:

$$\boldsymbol{E}^{\mathbb{Q}}(X) \approx \frac{1}{n}\sum_{k=0}^{n-1} X(\omega_k)$$

However, the calculation of conditional expectations $\boldsymbol{E}^{\mathbb{Q}}(X \mid \mathcal{F}_{T_i})$ is more involved, since the Monte Carlo simulation does not provide a discretisation of the filtration $\{\mathcal{F}_t\}$. For the estimation of conditional expectations, numerical approximations such as least-squares regressions can be used (see Fries 2007).

Let z denote a given model parameter used in the generation of the Monte Carlo simulation. Given a valuation algorithm that calculates the unconditional expectation $\boldsymbol{E}^{\mathbb{Q}}(V)$ of a random variable V, where the calculation of V involves one or more conditional expectations, we consider the calculation of the derivative $(\mathrm{d}/\mathrm{d}z)\boldsymbol{E}^{\mathbb{Q}}(V)$.

American Monte Carlo and Bermudan option valuation

We shall now give a brief definition of the backward algorithm. For more details, see Fries (2007); we use similar notation here.

Let $0 = T_0 < T_1 < \cdots < T_n$ denote a given time discretisation. Let V_i, $i = 1, \ldots, n$, denote the time-T_i numeraire relative values of given underlyings. Here, V_i are $\mathcal{F}_{T_i}$-measurable random variables. Then, let U_i be defined as follows:

$$\begin{aligned} U_{n+1} &:= 0 \\ U_i &:= B_i(\boldsymbol{E}^{\mathbb{Q}}(U_{i+1} \mid \mathcal{F}_{T_i}), U_{i+1}, V_i) \end{aligned} \tag{3}$$

where B_i is an arbitrary function.

Bermudan option valuation. For a Bermudan option, B_i is given by the optimal exercise criterion, ie, $B_i(x, u, v) := G(x - v, u, v)$, with:

$$G(y, u, v) := \begin{cases} u & \text{if } y > 0 \\ v & \text{else} \end{cases}$$

This defines a backward induction $i = n, n-1, \ldots, 1$ for U_i. For a Bermudan option, the unconditional expectation:

$$\boldsymbol{E}^{\mathbb{Q}}(U_1)$$

is the risk neutral (numeraire relative) value of the Bermudan option with exercise dates $T_1 < \cdots < T_n$ and exercise values V_i.

Bermudan digital option valuation. For a Bermudan option, exercise is optimal. This allows us to use a popular trick: ignoring the first-order derivative of $G(y, u, v)$ with respect to y. Hence, for a Bermudan option it is not necessary to differentiate B_i with respect to x (see Piterbarg 2004).

However, this is just a special property of the function G and holds for first-order derivatives only. For example, the trick cannot be applied to a Bermudan digital option, ie, $B_i(x, u, v) := H(x - v, u, v)$ with:

$$H(y, u, v) := \begin{cases} u & \text{if } y > 0 \\ 1 & \text{else} \end{cases}$$

We will use this product as a test case of the methodology in the numerical results.

[4] *It is actually sufficient that there is an outer operator, which is an expectation or a conditional expectation on a coarser filtration.*

REMARK 1 The Bermudan digital option shows a dependency on the first-order derivative with respect to the indicator condition y, since the expectation of the two outcomes u and 1 differs, ie, we have a jump at $y = 0$. For the classic Bermudan option, the expectations of u and v agree at $y = 0$.

The (automatic) differentiation of a jump may appear to be an issue, but it is possible to solve this elegantly in a stochastic automatic differentiation (Fries 2017), as we will see later.

Automatic differentiation of the American Monte Carlo backward algorithm

Let z denote an arbitrary model parameter. We assume the underlying values V_i depend on z.[5]

We now differentiate the backward algorithm (3). For $i = 1, \ldots, n$, we have:

$$\frac{\mathrm{d}}{\mathrm{d}z} U_i = \frac{\mathrm{d}}{\mathrm{d}z} B_i(\boldsymbol{E}^{\mathbb{Q}}(U_{i+1} \mid \mathcal{F}_{T_i}), U_{i+1}, V_i)$$

Applying the chain rule to $B_i(x, u, v)$ with $x = \boldsymbol{E}^{\mathbb{Q}}(U_{i+1} \mid \mathcal{F}_{T_i})$, $u = U_{i+1}$, $v = V_i$, and writing:

$$B_i = B_i(\boldsymbol{E}^{\mathbb{Q}}(U_{i+1} \mid \mathcal{F}_{T_i}), U_{i+1}, V_i)$$

to avoid the lengthy argument list, we get:

$$\frac{\mathrm{d}}{\mathrm{d}z} U_i = \frac{\mathrm{d}B_i}{\mathrm{d}x}\left(\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{E}^{\mathbb{Q}}(U_{i+1} \,\Big|\, \mathcal{F}_{T_i})\right) + \frac{\mathrm{d}B_i}{\mathrm{d}u}\frac{\mathrm{d}U_{i+1}}{\mathrm{d}z} + \frac{\mathrm{d}B_i}{\mathrm{d}v}\frac{\mathrm{d}V_i}{\mathrm{d}z}$$

and using $(\mathrm{d}/\mathrm{d}z)\boldsymbol{E}^{\mathbb{Q}}(U_i) = \boldsymbol{E}^{\mathbb{Q}}((\mathrm{d}/\mathrm{d}z)U_i)$:

$$\frac{\mathrm{d}}{\mathrm{d}z} U_i = \frac{\mathrm{d}B_i}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{i+1}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_i}\right) + \frac{\mathrm{d}B_i}{\mathrm{d}u}\frac{\mathrm{d}U_{i+1}}{\mathrm{d}z} + \frac{\mathrm{d}B_i}{\mathrm{d}v}\frac{\mathrm{d}V_i}{\mathrm{d}z} \tag{4}$$

Forward differentiation. Applying this relation iteratively (plugging the expression $\mathrm{d}U_{j+1}/\mathrm{d}z$ into the equation $\mathrm{d}U_i/\mathrm{d}z$ for $j = i, \ldots, n-1$) gives:

$$\frac{\mathrm{d}}{\mathrm{d}z} U_i = \sum_{j=i}^{n}\left(\left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{j+1}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_j}\right) + \left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}v}\frac{\mathrm{d}V_j}{\mathrm{d}z}\right) \tag{5}$$

Indeed, plugging (4) for $i+1$ into the right-hand side of (4) for i, we get:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z} U_i &= \frac{\mathrm{d}B_i}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{i+1}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_i}\right) + \frac{\mathrm{d}B_i}{\mathrm{d}v}\frac{\mathrm{d}V_i}{\mathrm{d}z} \\ &\quad + \frac{\mathrm{d}B_i}{\mathrm{d}u}\left[\frac{\mathrm{d}B_{i+1}}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{i+2}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_{i+1}}\right) + \frac{\mathrm{d}B_{i+1}}{\mathrm{d}u}\frac{\mathrm{d}U_{i+2}}{\mathrm{d}z} + \frac{\mathrm{d}B_{i+1}}{\mathrm{d}v}\frac{\mathrm{d}V_{i+1}}{\mathrm{d}z}\right] \\ &= \sum_{j=i}^{i+1}\left(\left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{j+1}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_j}\right) + \left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}v}\frac{\mathrm{d}V_j}{\mathrm{d}z}\right) \\ &\quad + \left(\prod_{k=i}^{i+1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}U_{i+2}}{\mathrm{d}z}\end{aligned}$$

Repeating this iteration, we get $i+1 \to n$ on the right-hand side. Using $\mathrm{d}U_{n+1}/\mathrm{d}z = 0$ gives (5).

To shorten our notation, let:

$$A_{i,j} = \left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}x}, \qquad C_{i,j} = \left(\prod_{k=i}^{j-1}\frac{\mathrm{d}B_k}{\mathrm{d}u}\right)\frac{\mathrm{d}B_j}{\mathrm{d}v}$$

so (5) becomes:

$$\frac{\mathrm{d}U_i}{\mathrm{d}z} = \sum_{j=i}^{n}\left(A_{i,j}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_{j+1}}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_j}\right) + C_{i,j}\frac{\mathrm{d}V_j}{\mathrm{d}z}\right) \tag{6}$$

This last equation would be natural in a forward (automatic) differentiation, since we calculate $0 = \mathrm{d}U_{n+1}/\mathrm{d}z$, $\mathrm{d}U_n/\mathrm{d}z$, $\mathrm{d}U_{n-1}/\mathrm{d}z, \ldots$, together with $\mathrm{d}V_j/\mathrm{d}z$, in a forward direction. Here, 'forward' refers to the order of operations in the algorithms, which run backwards over the indexes j.

Backward differentiation. Using backward (adjoint) automatic differentiation to calculate the derivative in (6) would require a mixture of backward propagation and the application of (6): we calculate $\mathrm{d}V_j/\mathrm{d}z$, followed by a forward application of (6). However, we can calculate the derivative in a single backward differentiation sweep. We start with (4) for $i = 1$. Since we are only interested in:

$$\frac{\mathrm{d}}{\mathrm{d}z}\boldsymbol{E}^{\mathbb{Q}}(U_1) = \boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}}{\mathrm{d}z}U_1\right)$$

taking the expectation in (4) we get:

$$\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}}{\mathrm{d}z}U_1\right) = \boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}B_1}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_2}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_1}\right) + \frac{\mathrm{d}B_1}{\mathrm{d}u}\frac{\mathrm{d}U_2}{\mathrm{d}z} + \frac{\mathrm{d}B_1}{\mathrm{d}v}\frac{\mathrm{d}V_1}{\mathrm{d}z}\right)$$

We may now use:

$$\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}B_1}{\mathrm{d}x}\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}U_2}{\mathrm{d}z} \,\Big|\, \mathcal{F}_{T_1}\right)\right) = \boldsymbol{E}^{\mathbb{Q}}\left(\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}B_1}{\mathrm{d}x} \,\Big|\, \mathcal{F}_{T_1}\right)\frac{\mathrm{d}U_2}{\mathrm{d}z}\right) \tag{7}$$

to get:

$$\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}}{\mathrm{d}z}U_1\right) = \boldsymbol{E}^{\mathbb{Q}}\left(\left(\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}B_1}{\mathrm{d}x} \,\Big|\, \mathcal{F}_{T_1}\right) + \frac{\mathrm{d}B_1}{\mathrm{d}u}\right)\frac{\mathrm{d}U_2}{\mathrm{d}z} + \frac{\mathrm{d}B_1}{\mathrm{d}v}\frac{\mathrm{d}V_1}{\mathrm{d}z}\right) \tag{8}$$

Plugging (4) into (8) and repeating the previous argument for $i = 2, \ldots, k-1$, we get the forward equation iteratively:

$$\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}}{\mathrm{d}z}U_1\right) = \boldsymbol{E}^{\mathbb{Q}}\left(A^*_{1,k}\frac{\mathrm{d}U_{k+1}}{\mathrm{d}z} + \sum_{j=1}^{k} C^*_{1,j}\frac{\mathrm{d}V_j}{\mathrm{d}z}\right)$$

where:

$$\begin{aligned}A^*_{1,i} &= \boldsymbol{E}^{\mathbb{Q}}\left(A^*_{1,i-1}\frac{\mathrm{d}B_i}{\mathrm{d}x} \,\Big|\, \mathcal{F}_{T_i}\right) + A^*_{1,i-1}\frac{\mathrm{d}B_i}{\mathrm{d}u} \\ C^*_{1,i} &= A^*_{1,i-1}\frac{\mathrm{d}B_i}{\mathrm{d}v} \\ A^*_{1,0} &= 1\end{aligned}$$

[5] *Think of z as an initial value of the interest rate curve (eg, calculating delta in a Libor market model (LMM)) or a volatility parameter (calculating vega).*

Using $k = n$, when $U_{n+1} = 0$, we have that:

$$\boldsymbol{E}^{\mathbb{Q}}\left(\frac{\mathrm{d}}{\mathrm{d}z}U_1\right) = \boldsymbol{E}^{\mathbb{Q}}\left(\sum_{j=1}^{n} C^*_{1,j}\frac{\mathrm{d}V_j}{\mathrm{d}z}\right)$$

The recursive definitions of $A^*_{1,i}$, $C^*_{1,i}$ have an intuitive interpretation in a backward (automatic) differentiation algorithm. In this algorithm, the conditional expectation operator on U_{i+1} is replaced by taking the conditional expectation of the adjoint differential.

Remark 2 The important improvement here (which highly simplifies the implementation) is that the calculation of the coefficients A^*, C^* does not involve the direct (automatic) differentiation of the conditional expectation operator. In addition, we can calculate the coefficients in a single backward differentiation sweep, due to the application of (7).

Automatic tracking of measurability. We can augment our implementation of random variables and operators by adding a property $\mathcal{T}(X)$ to a random variable X, such that X is $\mathcal{F}_s$-measurable for $s \geqslant \mathcal{T}(X)$. We set $\mathcal{T}(W(t)) := t$ for the Brownian driver W of our model and $\mathcal{T}(c) := -\infty$ for deterministic random variables c. For any operator $Z = f(X, Y)$, we set $\mathcal{T}(Z) := \max(\mathcal{T}(X), \mathcal{T}(Y))$.

Then, $\mathcal{T}$ enables the optimisation:

$$\boldsymbol{E}^{\mathbb{Q}}(X \mid \mathcal{F}_t) = X \quad \text{if } \mathcal{T}(X) \leqslant t$$

Differentiation of the indicator function

For a Bermudan digital option, the differentiation of B also contains the differentiation of the indicator function. The differentiation of the indicator function also appears in other products, and it is even relevant for the valuation of a Bermudan option, provided the estimation of the exercise boundary is not optimal. Hence, we have to consider:

$$\frac{\partial}{\partial X}\mathbf{1}(X > 0)$$

where:

$$\mathbf{1}(X > 0)(\omega) := \begin{cases} 1 & \text{for } X(\omega) > 0 \\ 0 & \text{else} \end{cases}$$

Automatic differentiation applied to algorithms involving indicator functions results in a linear combination of differentiations of those indicator functions. Hence, we have to evaluate expressions of the form:

$$A\frac{\partial}{\partial X}\mathbf{1}(X > 0)$$

where A is a linear operator (the adjoint differential).

If we are only interested in the expectation of the final result, it is sufficient to consider:

$$\boldsymbol{E}\left(A\frac{\partial}{\partial X}\mathbf{1}(X > 0)\right)$$

This evaluates to:

$$\boldsymbol{E}\left(A\frac{\partial}{\partial X}\mathbf{1}(X > 0)\right) = \boldsymbol{E}(A \mid \{X = 0\})$$

which can be approximated by

$$\boldsymbol{E}\left(A\frac{\partial}{\partial X}\mathbf{1}(X > 0)\right) \approx \boldsymbol{E}\left(A\frac{1}{2\delta}\mathbf{1}(|X| < \delta)\right) \tag{9}$$

In other words, if we are only interested in the expectation of the final result, we can approximate:

$$\frac{\partial}{\partial X}\mathbf{1}(X > 0) \approx \frac{1}{2\delta}\mathbf{1}(|X| < \delta)$$

This approximation has an intuitive interpretation. First, we may observe this approximation agrees with the result of a so-called payoff smoothing, where the indicator function is approximated by a call spread:

$$\mathbf{1}(X > 0)[\omega] \approx \begin{cases} \dfrac{X(\omega) + \delta}{2\delta} & \text{for } |X(\omega)| < \delta \\ 0 & \text{else} \end{cases}$$

More strikingly, the approximation is just a central finite-difference approximation of the derivative. It is:

$$\frac{\partial}{\partial X}\mathbf{1}(X > 0) \approx \frac{\mathbf{1}(X + \delta > 0) - \mathbf{1}(X - \delta > 0)}{2\delta}$$

Here, we have replaced the automatic differentiation by a (local) finite-difference approximation.

Since the implementation in Fries (2017) and at http://bit.ly/2FD86bo has access to the full random variable X, we can achieve an important improvement in the numerical algorithm: we can choose the δ shift that is appropriate for the random variable under consideration.

For example, we can choose the size of the bin δ in (9) as a multiple of the standard deviation of X:

$$\boldsymbol{E}\left(A\frac{\partial}{\partial X}\mathbf{1}(X > 0)\right) \approx \boldsymbol{E}\left(A\frac{1}{2\delta}\mathbf{1}(|X| < \delta)\right)$$

$$2\delta = \varepsilon(\boldsymbol{E}(X^2))^{1/2} \tag{10}$$

which essentially determines the number of paths used to approximate the conditional expectation (9) by a binning.[6] The effect of ε (or δ) is illustrated in our numerical results.

Our numerical results show there is an important advantage in applying the approximation (10) on a per-operator basis with an appropriate bin size δ (see figures 3 and 4).

Higher-order sensitivities

Higher-order sensitivities are made possible by applying the automatic differentiation to the lower-order automatic differentiation. This is immediately available in the implementation at http://bit.ly/2FD86bo (see the test cases there for examples).

The results presented here remain valid for higher-order sensitivities, since the automatic differentiation results in admissible operators, ie, operators on which the method is applicable. For example, the first-order differentiation of the indicator function results in a conditional expectation.

Numerical results

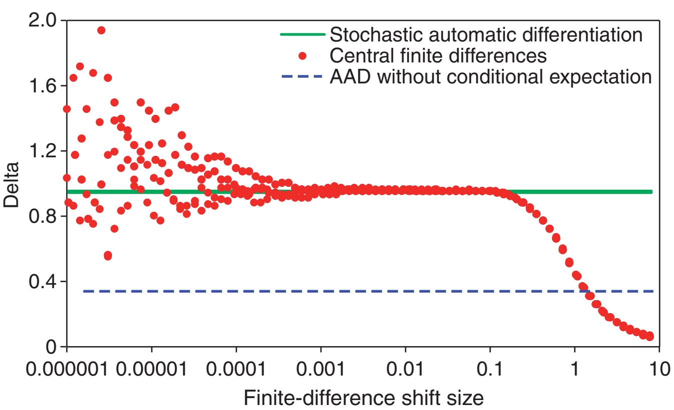
AAD delta of a Bermudan digital option. As a first test case, we calculate the delta of a Bermudan digital option paying:

1 if $S(T_i) - K_i > \tilde{U}(T_i)$ in T_i and if no payout occurred before

0 otherwise

[6] *Assuming a normally distributed X, a value of $\varepsilon = 0.2$ would result in approximately 8% of paths being used for the estimation of the conditional expectation (± 0.1 standard deviation).*

2 Delta of a Bermudan digital option using finite differences (red, different finite-difference shift sizes) and stochastic AAD (green)



where $\tilde{U}(T)$ is the time T value of the future payoffs for $T_1, \dots, T_n$. Note that $\tilde{U}(T_n) = 0$, so the last payment is a digital option.

This product is an ideal test case: the valuation of $\tilde{U}(T_i)$ is a conditional expectation. In addition, the conditional expectation only appears in the indicator function. The delta of a digital option payoff is only driven by the movement of the indicator function, since:

$$\frac{\mathrm{d}}{\mathrm{d}S_0}\boldsymbol{E}(\mathbf{1}(f(S(T)) > 0)) = \phi(f^{-1}(0))\frac{\mathrm{d}f(S)}{\mathrm{d}S_0}$$

where ϕ is the probability density of S and $\mathbf{1}(\cdot)$ is the indicator function (see Fries 2007). Hence, keeping the exercise boundary fixed would result in a delta of 0.

We calculate the delta of a Bermudan digital option under a Black-Scholes model ($S_0 = 1.0$, $r = 0.05$, $\sigma = 0.30$) using one million Monte Carlo paths. We consider an option with $T_1 = 1$, $T_2 = 2$, $T_3 = 3$, $T_4 = 4$, and $K_1 = 0.5$, $K_2 = 0.6$, $K_3 = 0.8$, $K_4 = 1.0$. The implementation of this test case is available at http://bit.ly/2FD86bo. The results are depicted in figure 2.

We depict the finite-difference approximation as red dots with different shift sizes on the x-axis. The finite-difference approximation was performed with different Monte Carlo seeds. We see the well-known effect that a finite-difference approximation of the derivative is biased for large shift sizes and unstable/unreliable for small shift sizes. We only see stable and unbiased results for shift sizes in the range 0.01–0.1.

We then depict the result of our method in green.[7] Our method reproduces the stable and unbiased result. Its Monte Carlo error is indistinguishable in the graph.

To illustrate that taking the conditional expectation of the adjoint differential is required, we repeat the calculation without this step. In blue, we depict the value of an automatic differentiation when the conditional expectation is omitted in the calculation of $A^*_{1,i}$ and reveal this would give a wrong result. Even worse, if we keep the exercise boundary fixed, the result is 0.

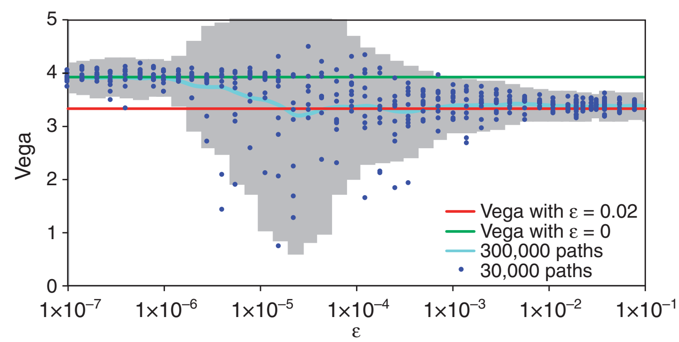
[7] *Since there is no shift size, there is no dependency on the shift size; thus, we get a horizontal line.*

A. Vega of a Bermudan swaption in the LMM

Algorithm	Evaluation	Derivative	Memory usage
Finite difference	2.0 s	25,600 s (7 hours)	1.5 GB
Efficient stochastic AAD	4.6 s	27 s	8.7 GB

Calculation times and memory usages for finite differences and backward algorithmic differentiation algorithms. LMM with 15,000 paths and 25,600 theoretical model vegas

3 Vega of a Bermudan swaption using AAD for different values of ε for the differentiation of the indicator function in (10)



The calculations were performed for 30,000 paths with different bin sizes ε (blue). We depict a mean (blue line) and standard deviation (grey). The green line marks the vega for $\varepsilon = 0$ (ignoring the differentiation of the indicator function). The red line marks the vega for $\varepsilon = 0.2$. Small values of epsilon result in unstable vegas, since only a few paths are used for the estimation. The value 0.2 can be interpreted statistically using 8% of the paths for the estimation of the derivative

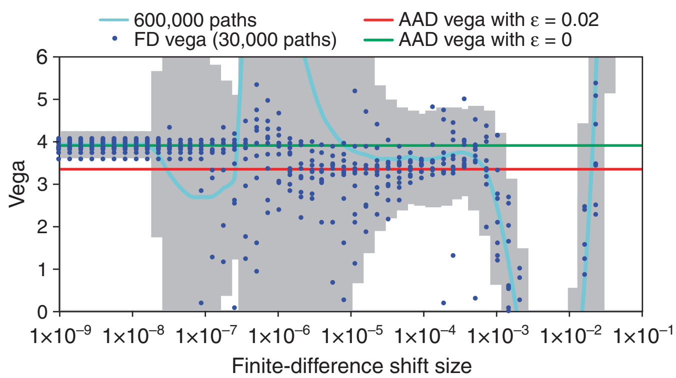
AAD vega of a Bermudan option (under LMM). As a second test case, we consider the AAD calculation of the vega of a Bermudan swaption (30Y in 40Y, with semi-annual exercise) under the LMM (40Y, semi-annual, 320 time steps, 15,000 paths). The model has 25,600 independent instantaneous volatility parameters, resulting in 12,640 effective vega sensitivities. The performance of the algorithm is given in table A.

To analyse the numerical stability, we depict the values of a parallel vega of the Bermudan swaption, where the conditional expectation is estimated by a least-squares regression. In theory, the differentiation of the conditional expectation can be ignored, since the Bermudan swaption valuation can assume optimal exercises (see Piterbarg 2004). However, the Monte Carlo error and the regression error will result in a non-optimal exercise (see Fries 2006), which results in a (slightly) different vega. Assuming optimal exercise will result in wrong hedge ratios, since an exercise will be effectively suboptimal due to the biased valuation of the future exercise dates. In addition, a comparison of the two different vegas (with and without ignoring differentiation of the indicator function) gives an indication of the quality of the conditional expectation estimator.

We may control the differentiation of the exercise boundary through the parameter ε in (10) (where $\varepsilon = 0$ is interpreted as ignoring the differentiation of the indicator function). The parameter ε thus enables us to investigate the optimality of each exercise boundary.

In figure 3, we show the AAD vega for different values of ε for the size of the bin estimating the derivative of the indicator function. Since ε determines the number of Monte Carlo paths used for sampling the differentiation of the indicator functions, we see a larger Monte Carlo error

4 Vega of a Bermudan swaption using finite differences (blue, different random number seeds and finite-difference shifts)



The red and green lines mark the values obtained by AAD, where the differentiation of the indicator function was performed using (10) with $\varepsilon = 0$, ignoring differentiation of the indicator function (green), and $\varepsilon = 0.2$ (red), resulting in two different values for vega. The patterns visible for finite-difference vegas are due to the same Monte Carlo seed being used for different shift sizes, generating a jump in vega once a Monte Carlo path crosses the exercise boundary

for very small ε. The value $\varepsilon = 0$ will switch off the differentiation of the indicator function and hence result in the AAD differentiation ignoring the differentiation of the indicator function.

In figure 4, we depict the classic, brute-force finite-difference approximation of the vega for different shift sizes and the stochastic AAD calculation of vega for $\varepsilon = 0$ (green) and $\varepsilon = 0.2$ (red). We see again that the finite-difference approximation ignores the differentiation of the exercise boundary for very small shifts. This is because no path crosses the exercise boundary under a small shift. For larger shifts, the finite-difference approximation includes the differentiation of the exercise boundary but results in a huge Monte Carlo variance.

It is almost impossible to obtain a reliable estimate for the Bermudan swaption vega by finite differences, which includes the effect of a suboptimal exercise due to differentiation of the indicator functions. AAD gives a reliable estimate (red). The finite-difference sample points show some convergence to the correct solution for a narrow range of shifts sized around 0.005.

A comparison of figures 4 and 3 highlights an advantage of the AAD differentiation of the indicator function: while a finite-difference shift on the parameter acts on all indicator functions (eg, all exercise dates) and an optimal or good choice of the finite-difference shift is unclear, the AAD differentiation allows for a per-operator determination of the size of the estimation bin.

Implementation design

The implementation can be performed with a minimum of code complexity and in few lines of code. The implementation contains an automatic tracking of measurability and a per-operator estimation of the derivative of the indicator functions. See http://bit.ly/2tBxDQK, http://bit.ly/2FD86bo and Fries (2017) for details.

Conclusion

We presented a modification of the backward automatic differentiation to apply a true adjoint algorithmic differentiation to algorithms that require the estimation of conditional expectations and use indicator functions. This algorithm avoids the differentiation of an approximation of the conditional expectations to improve accuracy and performance. For indicator functions, it enables us to perform a per-operator calculation of the differentiation, allowing accurate treatment of individual exercise boundaries. ■

Christian Fries is head of model development and methodology at DZ Bank in Frankfurt and a professor of mathematical finance at LMU Munich.
Email: email@christian-fries.de.

REFERENCES

Andreasen J, 2014
CVA on an iPad Mini
Global Derivatives, ICBI, Amsterdam, The Netherlands

Antonov A, 2017
Algorithmic differentiation for callable exotics
Working Paper, April 4, SSRN, available at http://ssrn.com/abstract=2839362

Antonov A, S Issakov, M Konikov, A McClelland and S Mechkov, 2018
Pathwise XVA Greeks for early-exercise products
Risk January, pages 74–79

Capriotti L and M Giles, 2012
Algorithmic differentiation: adjoint Greeks made easy
Risk September, pages 96–102

Capriotti L, Y Jiang and A Macrina, 2017
AAD and least squares Monte Carlo: fast Bermudan-style options and XVA Greeks
Algorithmic Finance 6(1–2), pages 35–49

Fries CP, 2006
Foresight bias and suboptimality correction in Monte-Carlo pricing of options with early exercise
In *Progress in Industrial Mathematics at ECMI 2006*, pages 645–649
Springer

Fries CP, 2007
Mathematical Finance: Theory, Modeling, Implementation
Wiley

Fries CP, 2017
Stochastic automatic differentiation: efficient implementation of automatic differentiation for Monte-Carlo simulations
Working Paper, June, SSRN, available at https://ssrn.com/abstract=2995695

Fries CP, 2019
Fast stochastic forward sensitivities in Monte-Carlo simulations using stochastic automatic differentiation
Journal of Computational Finance 22(4), forthcoming

Giles M and P Glasserman, 2006
Smoking adjoints: fast Monte Carlo Greeks
Risk January, pages 92–96

Piterbarg V, 2004
Computing deltas of callable LIBOR exotics in forward LIBOR models
Journal of Computational Finance 7, pages 107–144