mni

All Signal, No Noise

MARKET ANALYSIS

BOE September APF Decision: What you need to know

Contents

Tim Davis, 4 July 2025

- We think that the market is looking at the wrong measure of reserves when discussing the BOE's September APF decision: we think that supply-led reserves are much more important than total reserves.
- Supply-led reserves will fall below the top of the preferred minimum range of reserves (PMRR) when taking into account QT for the current period (to Sep25) and TFSME repayments (to Oct 25).
- Even with just passive QT from Oct25-Sep26, supply-led reserves will fall to the middle of the PMRR range.
- We outline the fiscal costs of QT (which are huge).
- At present we think the best decision for QT in the year ahead would be to continue active sales at broadly the same pace as in the current year.
- We also set out in our appendices the main components of the Bank's balance sheet and how the repo operations work

We are having increasing dialogue with clients regarding potential changes to QT for the 25/26 year (which is due to begin in October). The decision on the target reduction in the APF will be made at the September MPC meeting and published alongside the Bank Rate decision on 18 September. However, the implementation decision is made by the Bank Executive rather than the MPC (although there is of course crossover between the internals on the MPC here). In order to understand the moving parts in the decision and their relevance we set out what are the drivers of supply and demand of reserves.

We think that the market should be looking at supply-led reserve balances (not total reserve balances) as the backdrop for the September APF decision

We think that part of the market discussion around whether the Bank will continue with active gilt sales from September onwards is focused on the wrong number. Most of the arguments we have seen compare the total level of reserve balances to the PMRR. However, we think that it is important to split these reserve balances between supply-led balances and demand-led balances. We classify demand-led balances as those which the market can control in the short/medium-term – primarily the STR and ILTR. We classify other reserves as supply-led: these are broadly made up of the Bank's loan to the APF plus TFSME holdings minus physical money in circulation (there are other items that come into the equation here but these are generally so small that they don't have an overall significant impact).

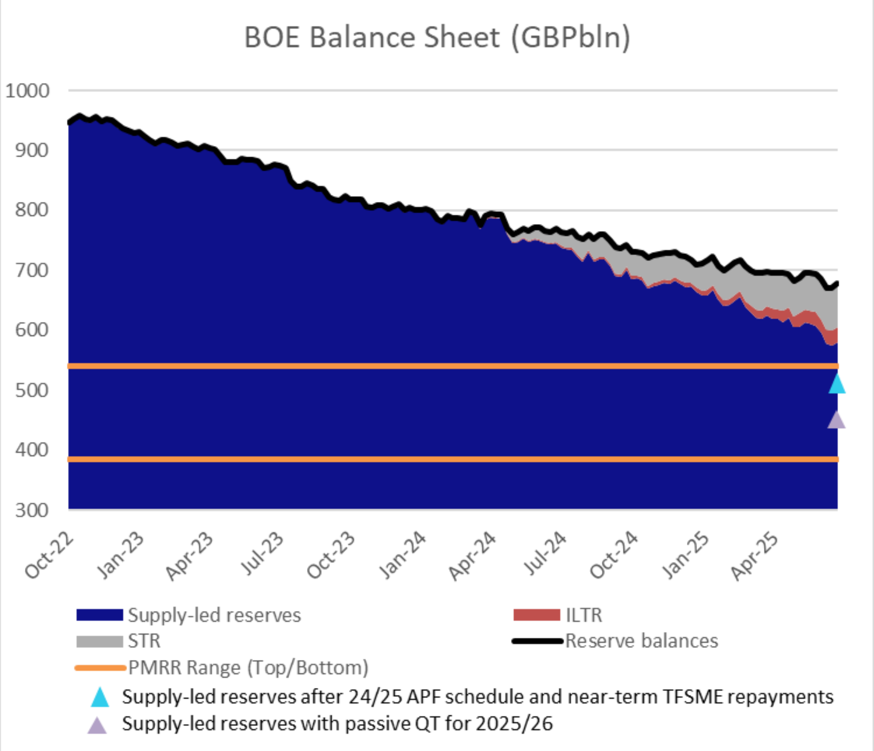
When breaking reserves down into these classifications, the size of the supply-led balance is by far the most important from the perspective of the Bank's decision on the size and scope of QT operations for the 25/26 year at the September MPC meeting (assuming that we don't see stresses in the banking system before then). We calculate these supply-led balances to equal GBP590bln at present (almost GBP100bln smaller than GBP678bln of total reserves).

Irrespective of the September decision there will be further reductions in the supply-led balances:

- The APF will reduce by a further GBP32bln in the current 24/25 year (ending September 2025). GBP28.3bln of this is from the 2.00% Sep-25 gilt that is due to mature on 7 September (GBP26.2bln nominal). There is also a final APF active gilt sales operation for each maturity bucket due in July.
- There is likely to be in excess of GBP40bln of further repayments of TFSME borrowings from the current level of GBP84bln to somewhere a bit north of GBP40bln by the end of October.

Together these will equate to supply-led balances reducing by more than GBP70bln between now and the end of October. This would reduce supply-led balances to around GBP510bln. We think this is hugely significant because if the upper limit for PMRR is GBP540bln, supply-led reserves will be below this level in less than four months time.

In addition, even if active gilt sales are stopped there would be another GBP49bln reduction in the year from October 2025 to September 2026 (GBP19.8bln of the 0.125% Jan-26 and GBP29.7bln of the 1.50% Jul-26). This means that by mid-2026 supply-led reserves will be below GBP460bln (with downside risks if we have overestimated the residual size of the TFSME past October 2025). GBP460bln is actually closer to the mid-point of the GBP385-540bln PMRR estimated range than the top end.



Source: MNI, Bloomberg Finance L.P., Bank of England
Note: Supply-led reserves are reserve balances minus "demand-led" ILTR / STR usage

The objectives for QT are important to bear in mind – we think two of these goals are broadly met

There are a few options for the Bank of England here, but before looking at these we will point to the communications from MPC members and senior members of the BOE Markets team. Ahead of the decision on the pace of balance sheet reduction last year, we were advised to look at the past for precedent. That is not a line that we have seen in any commentary this time around. The communication has been that this year is different and there are more factors to take into account.

But also from a policy perspective, it is important to look at the rationale for QT. In simple terms these were:

1. Reduce excess reserves and move back to an ample reserves system.
2. Reduce interest rate exposure from the BOE.
3. Bring reserves down to a level where QE or above schemes like TFSME could potentially be used as a policy tool once again if needed (with potency).
4. Not impact market conditions adversely and maintain financial stability.

We would argue that objective 1 has almost been met now – and probably will have been by the end of the current APF year and TFSME repayment period (and is very likely to have been purely through the passive roll off next year). Objective 3 is arguably very closely related to objective 1, too.

This leaves objective 2 (which is still very much a work in progress) and objective 4. Objective 4 can be seen in action both through the delaying of active gilt sales at the beginning of the QT period (with the Truss government/LDI crisis) and more recently through the rescheduling of the long-dated APF sales operation being replaced with a short-dated APF operation.

The costs of QT are huge – no wonder there is political pressure to slow it

We note that active QT is not a costless policy either. Not only does QT lead to private holdings of gilts increasing through both redemptions (as new private investors will need to buy the gilts from the DMO that the BOE did not roll over) and through the sales process (which increases net supply of gilts to the market) but it also leads to higher gilt issuance through the Treasury paying money to the APF to make up for the losses it has made. See Appendix 2 for more on the accounting here.

These losses are not insignificant either. Through the active sales process so far GBP69.7bln has been raised but this has reduced the size of the APF balance sheet by GBP104.1bln – a loss of GBP32.4bln since November 2022. Furthermore between March 2022 and September 2025 (the end of the current APF year) passive redemptions will have booked a loss of GBP15.3bln. So, losses for the APF since QT began from sales and redemptions have already reached GBP47.7bln – a sum that has been financed by the Treasury through higher gilts sales via the DMO.

Putting that in broad terms we are looking at roughly GBP15bln per year. With Bank Rate at 4% there is also a GBP10.8bln per year net interest income cost (which would drop to GBP5.2bln if Bank Rate fell to 3% and would only fall into a small surplus if Bank Rate fell to 2% - well below neutral).

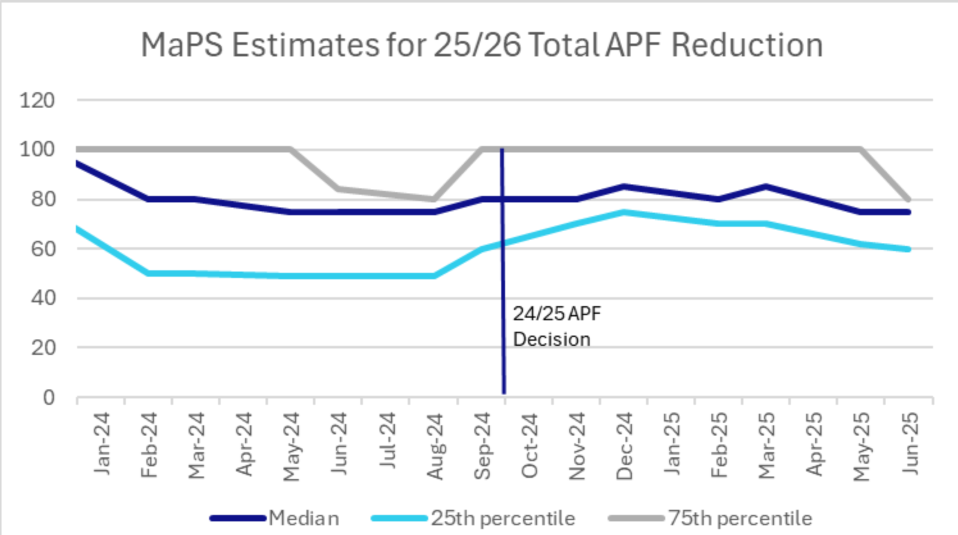
With the intention of the government's welfare bill having been to save GBP5bln, but with all the concessions being knocked down to around GBP2.0-2.5bln likely savings, this really puts the scale of these numbers into perspective from a political perspective.

So this leaves us in a situation where there is huge political pressure to stop or slow active gilt sales and with reducing interest rate exposure the only real benefit to the BOE, the positives of continuing the APF at scale are less clear.

Consensus estimates for balance sheet reduction

The median estimate for balance sheet reduction in the year from October 2025 through September 2026 in the MaPS survey has been relatively steady and remains around the GBP75bln mark (which would be roughly doubling the active QT pace that we have seen this year from GBP13bln to GBP26bln).

However, the 75th percentile estimate dropped from GBP100bln (where it had been static since last year's announcement) to GBP80bln in the June survey. The 25th percentile has also been gradually falling lower since last year's decision, too. So market estimates are beginning to fall to some extent (even if this isn't visible in the median estimate yet).



Source: MNI, Bank of England Market Participants Survey

Main options for the BOE's September 2025 APF decision:

The decision on the target reduction in the APF will be made at the September MPC meeting and published alongside the Bank Rate decision on 18 September. However, the implementation decision is made by the Bank Executive rather than the MPC (although there is of course crossover between the internals on the MPC here). The latter will be published in a Market Notice alongside the decision.

Option 1: Increase the pace of active QT while maintaining the current maturity buckets. This seems to be the consensus view at present (around GBP26bln per year), but we would argue is the least efficient. As we note above objective 4 of not having an adverse market impact would be at risk here in our view. The BOE has already had to postpone one long-dated operation in Q2-25 and doubling the volume of long-dated sales would not be prudent in our view.

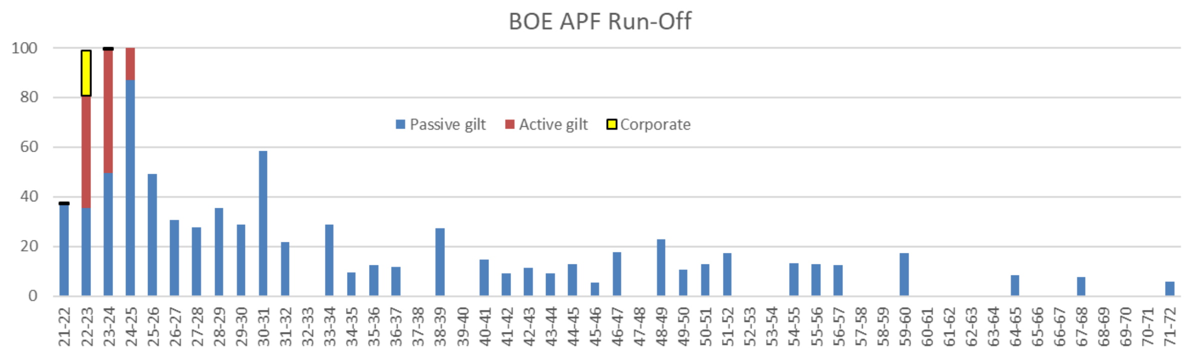
Option 2: Increase the pace of active QT in line with consensus expectations (double pace of this year) but with a skew towards shorter-dated gilts. The argument here is that the Bank of England has in the past argued that it has conducted its operations equally between maturity buckets so as to be curve neutral. It is quite evident that selling shorter-dated gilts has much less of a market / curve impact than selling at the long-end. It is not clear if there has been a permanent structural change to the market here, but the answer to that is unlikely to be resolved before the Bank makes its decision. We would favour some long-dated gilt sales operations to continue – possibly even at a slower pace than seen this year. This would have the benefit of off-the-run gilts at a range of maturities being supplied on a regular basis – which is something the market values and has been introduced to the DMO's sales programme in an alternative way via the introduction of PGTs (programmatic gilt tenders). This could be a halfway house – helping the BOE reduce interest rate exposure but also avoiding negative market connotations. It also has the advantage that shorter-dated operations crystalise smaller losses and hence have less cost to the Treasury.

Option 3: Keep the pace of active QT in line with the 24/25 pace (GBP13bln per year). This would be broadly in line with the 25th percentile response for the MaPS survey which has looked for GBP60-62bln over the last couple of surveys. This has the advantage of active sales being small enough that if market conditions appear to not favour selling at the long-end in any particular quarter, that operation could be replaced with another short/medium-dated operation without negatively impacting the market for those buckets. As we noted above in option 2, this also allows a decent selection of off-the-run gilts to be supplied to the market on a regular basis, too.

Option 4: Stop active gilt sales. We would not completely rule this out. As we argued above, supply-led reserves will be towards the middle of the estimated PMRR range by the summer 2026 even without active gilt sales. As we approach that point the benefits are solely to do with interest rate risk being held by the Bank.

We favour option 3 with a change to maturity buckets

If the decision was down to us, we would favour option 3, but with a change to the maturity buckets to re-align with the DMO. That would see short remaining at 3-7 years, medium to 7-15 years (rather than BOE's 7-20 years) and long 15+ years (rather than 20+ years). This gives ample flexibility, helps provide the market with a steady supply of off-the-run gilts across a variety of maturities and is able to be quickly adapted if market conditions were to sour at all. Aligning the buckets with the DMO also allows 15-20 year gilts to be sold back to the market if there is more demand in that area than in the 20+ year area.

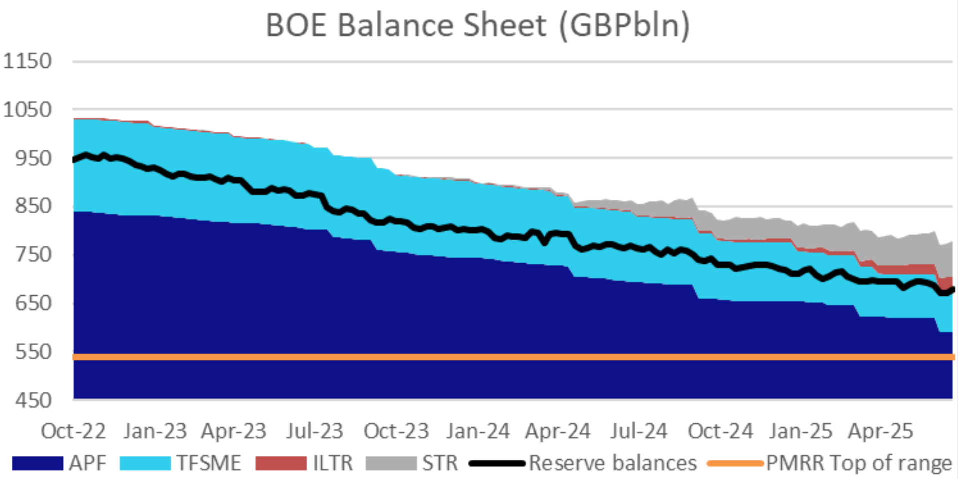


Source: MNI, Bank of England
Note: Passive gilt run-off will be reduced from 3+ years out due to active sales. Years run from October to September (e.g. 24-25 is from October 2024 to September 2025).

Appendix 1: What are the main items on the Bank of England's balance sheet

The main drivers of supply of reserves are the stock of gilts held in the APF, lending to commercial banks via the TFSME (Term Funding for SMEs) scheme, the Short Term Repo (STR) and Indexed Long-Term Repo (ILTR) facility. Broadly on the other side of the ledger are reserve balances and banknotes and coinage (we will ignore the cash in circulation in this discussion).

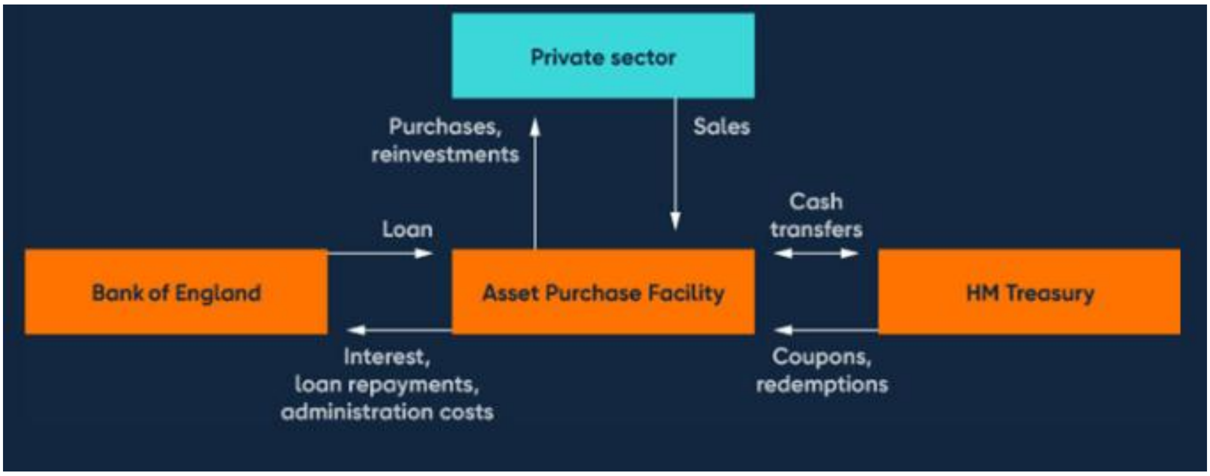
Demand for reserves has increased since the global financial crisis, partly due to regulatory reasons and partly due to banks' wanting to hold higher levels of reserves for market signalling reasons. However, it is widely believed that there are still excess reserves in the system at present relative to demand. Reserve balances currently stand at GBP678bln after moving as low as GBP670bln in recent weeks (the peak in January 2022 was GBP979bln – so excess reserves have already moved around GBP300bln below their peak). The Bank's survey shows that the expected preferred minimum range of reserves (PMRR) is GBP385-540bln. As we transition towards this level, reserve balances will transition from being supply-driven to demand-driven – with repo to take an increasingly prominent role.



Source: MNI, Bloomberg Finance L.P., Bank of England
Note: Reserve balances is broadly these 4 components minus physical currency

Appendix 2: How the accounting of the APF works and why QT increases gilt issuance

The Bank's APF scheme is technically a separate entity to the main Bank of England account and is financed by a loan from the Bank to the APF facility (which is fully indemnified by the Treasury). The loan from the BOE to the APF increased with each purchase of gilts (or corporate bonds) to facilitate the purchase i.e. if there was a GBP1.0bln gilt purchase operation, GBP1.0bln was lent from the BOE to the APF. The loan from the BOE to the APF is remunerated at Bank Rate, which was of course very cheap at the time of Bank Rate being very low but resulted in much higher interest payments on the loan from the APF as Bank Rate rose.



Source: Bank of England

Gilts bought for QE purchases on the APF balance sheet are not marked to market and are held at their purchase price until either being sold or until maturity. The vast majority of gilts held by the APF were bought at a price above par, so either a sale (at a lower price) or the gilt maturing at par will mean that the APF crystalises a loss when the gilt is disposed of. The APF also receives income from the coupon payments the Treasury pays for the gilts that it holds.

Although not originally set up this way, it was relatively quickly changed so that rather than profits / losses of the APF remaining on the APF's balance sheet that there would be a quarterly transfer of cash between the APF and the Treasury. Initially and through the QE phase, this meant there would be quarterly payments from the APF to the Treasury as the coupon income from the gilts held far exceeded payments to the BOE (for the loan interest at Bank Rate and the running costs of the APF).

However, as Bank Rate rose above the average coupon rate of the APF's gilt portfolio, rather than the APF making a profit and paying this to the Treasury, the APF began to make a loss and the Treasury would instead make a quarterly payment to the APF.

The losses are not just due to the net interest income flipping, but as we note above almost every time a gilt matures (i.e. passive QT) or is sold (active QT) this generally results in a loss for the APF – which results in even higher payments from the Treasury to the APF.

This of course has the impact that the higher the volumes of active sales, the higher the near-term cost to the Treasury. Note that if yields were to remain at current levels, the costs over time shouldn't theoretically be hugely dissimilar – but the losses would be spread over a number of years rather than front-loaded.

Also note that the higher duration the bond, the larger the near-term loss. And by definition, longer maturities and lower coupons are higher duration.

So there are two controversial points here that have been picked up by a number of political parties and is consistently asked at the Treasury Select Committee and Lords Economic Affairs Committee when members of the Bank of England testify. The first is whether active sales should continue at all and whether it is "right" from a wider policy perspective to have the Treasury needing to increase borrowing of gilts in the near-term in order to transfer this cash to the APF. And second, if active sales are to continue, should sales of high duration gilts be limited to some extent. We discuss this above in our options for September section.

Appendix 3: Term Funding for SMEs scheme

The TFSME was set up during Covid (and financed via reserves) to incentivise banks to keep lending to SMEs by allowing them to borrow on a four-year term at rates very close to bank rate with eligible participants allowed to borrow at least 10% of their stock of real economy lending with additional borrowing allowances generated at £1:£1 for lending to households and £5:£1 for lending to SMEs.

The scheme was open for applications between 15 April 2020 and 31 October 2021 with takeup peaking at GBP192.9bln. There are no penalties to repaying funds early and at present GBP84.2bln remains outstanding. Some of this lending (which backed the bounceback loan scheme) has been extended to between 6-10 years – we don't know exact numbers but expect at least another GBP40bln or so will be repaid between now and the end of October.

Appendix 4: Short Term Repo facility

The STR is one of the two weekly demand-led pillars of the Bank of England's repo system. An auction is held on Thursday morning each week (afternoon on MPC decision days) and allows participants to borrow reserves for a one-week period (or 2 weeks over Christmas) in exchange for the highest quality collateral (technically referred to as Level A). The operation is full allotment at a fixed price which is equal to Bank Rate.

The STR was set up in October 2022 but usage only really began to pick up around April 2024 and there has been a gradual increase in usage since then with the latest operation seeing take up of GBP74.2bln.

The increase in takeup in here is due to a few different factors. Most market participants (in a view held by the BOE) don't expect the increased take up to indicate that we have reached PMRR, but as excess reserves have become a little less prevalent, increased use of the facility has become normalised. And with GC repo rates rising a little above Bank Rate there has been some added incentive to use the STR rather than other repo facilities in the market.

There are two elements that are expected to limit STR usage (and push towards the ILTR) in the medium-term. First, the requirement for the high quality collateral which some market participants won't have as readily available. And second, the fact that the term is only one-week.

Appendix 5: Indexed Long-Term Repo operations (ILTR)

The other pillar of the demand-led repo system is the ILTR. This is a weekly operation held every Tuesday in which a maximum of GBP35bln of reserves at each operation (up from GBP 25bln prior to 11 June 2025) are able to be exchanged for a broader range of collateral than just the Level A highest quality collateral that is eligible for STR operations.

Also unlike the STR operation, the ILTR operations are both variable rate and variable in size (although they are still uniform price auctions within each collateral level). Level A collateral can currently be bid at Bank Rate, but the minimum spread from November will increase to 3bps above Bank Rate. For Level B collateral the minimum spread will be 5bps above Bank Rate and for Level C the minimum spread will be 15bp.

The June 2025 recalibration of the ILTR saw the quantity of reserves available at the minimum clearing spreads rise to GBP8bln per auction from GBP5bln per auction (GBP4bln for Level A, GBP0.8bln for Level B and GBP3.2bln for Level C).

We note that since the ILTR was set up in 2014 (albeit monthly rather than weekly originally) the vast majority of usage has been at the minimum spread. The most recent ILTR operation which did not see usage cleared at the minimum spread was on 24 March 2020 – the day after the first Covid lockdown was announced.

It is envisaged by the Bank of England that when reserves broadly reach PMRR levels that STR usage will broadly level out and as the APF is wound down further usage of the ILTR increases in its place. So eventually it is foreseen that ILTR usage will be higher than STR usage.

Unauthorized disclosure, publication, redistribution or further dissemination of this information may result in criminal prosecution or other severe penalties. Any such authorization requires the prior written consent of Market News International. Redistribution of this information, even at the instruction of your employer, may result in personal liability or criminal action unless such redistribution is expressly authorized in writing by Market News International. Violators will be prosecuted. This information has been obtained or derived from sources believed to be reliable, but we make no representation or warranty as to its accuracy or completeness. This is not an offer or solicitation of an offer to buy/sell. Copyright © 2025 Market News International, Inc. All rights reserved.