I am grateful to Saba Alam, Nicholas Butt, Amber Evans, Rafael Kinston, Nick McLaren, Dave Ramsden, James Rapa, Rhys Phillips, Tom Smith and Tim Taylor for their assistance, guidance and advice in preparing this speech.

All speeches are available online at www.bankofengland.co.uk/publications/Pages/speeches/default.aspx
Introduction

For such an apparently arcane topic, the size and composition of central bank balance sheets can provoke passionate reactions.

‘Too large’, and central banks may find themselves accused of usurping the role of financial markets, harming innovation and inducing imprudent behaviour; fuzzying the boundary between monetary and fiscal policy, providing a ‘dangerous temptation for … the political class’; or giving unmerited financial rewards to reserves holders.\(^2\)

‘Too small’, and central banks may be criticised for being asleep at the wheel\(^3\) at times of crisis; failing to play their part in ensuring an adequate supply of risk-free assets in the economy to maintain financial stability during peacetime\(^4\); or hampering the effectiveness of monetary policy transmission.\(^5\)

Chart 1: Central bank balance sheets compared

The Bank of England has found itself on both ends of this debate in the past 10-15 years. Before the financial crisis, our balance sheet was modest, at 4% of GDP. Since then, and in direct response to the crisis, that figure has risen to around 30%: a more than seven-fold increase. Other regions have seen similar, or in some cases much larger, expansions (Chart 1). But by the UK’s own standards this is truly exceptional. You have to go back to the end of World War 2 or the early 18\(^{th}\) century to find anything even approaching the same highs (Chart 2).

---

\(^1\) Kevin Warsh, quoted in the FT, 7.5.17: [https://www.ft.com/content/45303028-3280-11e7-bce4-9023f8c0fd2e](https://www.ft.com/content/45303028-3280-11e7-bce4-9023f8c0fd2e)

\(^2\) See for instance [https://www.cato.org/blog/strange-official-economics-interest-excess-reserves](https://www.cato.org/blog/strange-official-economics-interest-excess-reserves)

\(^3\) [https://www.thetimes.co.uk/article/we-were-not-asleep-at-the-wheel-bank-of-england-managers-tell-angry-mps-6pc7bnr56lf](https://www.thetimes.co.uk/article/we-were-not-asleep-at-the-wheel-bank-of-england-managers-tell-angry-mps-6pc7bnr56lf)


\(^5\) [https://www.brookings.edu/blog/ben-bernanke/2016/09/02/should-the-fed-keep-its-balance-sheet-large/](https://www.brookings.edu/blog/ben-bernanke/2016/09/02/should-the-fed-keep-its-balance-sheet-large/)
Where is this all going? At one level the answer to that is simple: for the foreseeable future, the path of the Bank’s balance sheet will be shaped by the decisions of the Monetary Policy Committee (MPC) in pursuit of the inflation target given to it by Government. Whether that involves further expansion of the balance sheet, or a return to less exceptional levels (so-called Quantitative Tightening or ‘QT’), we on the Bank executive need to be ready to implement those decisions.

My focus today is squarely on the latter: how we plan to implement monetary policy during QT, and the likely size of our balance sheet in the medium term. Of course, market expectations don’t point to QT starting any time soon. But those views can change rapidly. And experience from the United States and elsewhere suggests that early, clear communication about the target end-state can help stabilise expectations and increase the effectiveness of normalisation when it comes. For that reason, we launched a Discussion Paper⁶ last year to expose our vision of the medium-term goal to challenge and feedback. We promised to share our findings from that work⁷ – and that’s what I aim to do today.

To jump to the punchline, I have four key messages. The first is that, judged by historical standards, big balance sheets are here to stay. That’s not a prediction that QE will never unwind: it will. But we have a bigger responsibility than we did to provide liquidity to the system, in good times and bad, and to a wider set of organisations, to maintain financial stability. And that’s not going away. Point two is that big doesn’t mean outsized – so the balance sheet will eventually shrink from where it is today. That’s something the Bank has been stressing for some time. But the Discussion Paper has allowed us to put a tighter range on that forecast, and suggests our liabilities probably only need to be half the size they are today to carry out our mission once QT is underway. Knowing that is progress — but point three is we can’t get too comfortable,

---

because the margin of uncertainty around that estimate is high. Neither we nor the firms who use our
liquidity really know what their demand will be when conditions normalise. And that uncertainty could pose
challenges as we exit, as the recent experience of QT in the United States has arguably shown. The final
message, therefore, is that we must have as our ultimate goal an end-state framework that can cope with
that ambiguity without shaking itself, and us, to bits. We think we have an approach that can meet that
challenge – but we need to keep testing that, with your help.


To develop these points, let me start with some basics. The Bank of England’s balance sheet is special for
one key reason – and that’s the nature of our liabilities. Central bank money – whether banknotes or central
bank ‘reserves’ (deposits held with us by financial institutions) – provides the ultimate means of settlement
for all sterling payments in the economy. And that gives our balance sheet a central role in supporting
monetary and financial stability:

- It supports financial stability by providing the reserves banks need to make payments to each other
  and help insure against bank runs;⁸ and

- It underpins monetary stability because the interest rate we pay on reserves – known as Bank Rate
  – influences interest rates throughout the economy; and we can use reserves to conduct less
  conventional monetary policy, include asset purchases and term lending.

This last point has of course been the primary driver of our balance sheet expansion over the past decade.
Three quarters of the Bank’s assets is in the form of a loan to the Asset Purchase Facility backing £435bn of
gilt holdings and £10bn of corporate bonds, while another £127bn has been lent to banks under the
Term Funding Scheme (Chart 3).⁹ A further £13bn of liquidity has been extended under the so-called
‘Index Linked Term Repo’ facility, part of the Sterling Monetary Framework (SMF).

Nearly all of that activity has been financed by an increase in central bank reserves (Chart 4).

---

⁸ Alongside meeting regulatory liquidity requirements, and deposit insurance.
⁹ https://www.bankofengland.co.uk/markets/funding-for-lending-and-other-market-operations
Although created to finance asset purchases and loans, this increase in reserves supply has also allowed us to (more than) meet a sharp pickup in the demand for reserves as a safe asset, underpinning financial stability. That higher demand comes from two main sources:

- First, the number of firms with access to the Bank’s balance sheet through the SMF has increased ten-fold, from 17 in 2005 to over 200 today.\(^\text{10}\) That follows a deliberate push on our part to expand access to smaller banks, CCPs and broker-dealers as part of ensuring the Bank is ‘open for business’ and responsive to changing market needs.

- Second, there has been a sharp rise in SMF members’ target liquidity holdings, reflecting both the material strengthening in post-crisis regulatory requirements, and firms’ wish to hold buffers well above those minima. To give a sense of the scale of this effect, major UK banks now hold more than £1 trillion of assets classed as liquid for regulatory purposes – more than four times higher than pre-crisis. A quarter of that is held as central bank reserves.

In addition to implementing unconventional monetary policy and meeting firms’ financial stability needs, the Bank’s balance sheet operations have also delivered effective interest rate control. The key to this is that there is an abundance of reserves, all remunerated at Bank Rate. The excess of reserves pushes down on overnight markets rates, but they are prevented from falling much below a floor of Bank Rate by the fact that banks can borrow reserves in the market and earn Bank Rate by depositing them at the Bank of England. A

key merit of this so-called ‘floor’ system is that it can keep overnight interest rates close to Bank Rate despite large changes in the supply of, or demand for, reserves (Chart 5). And rates have indeed been closer to Bank Rate on average than at any point in the past twenty years (Chart 6). The regime in place prior to QE, known as ‘reserves averaging’, was also successful at controlling interest rates when the stock of reserves was relatively stable. But it did not cope as well when demand for reserves spiked and we had to inject large quantities of reserves for liquidity insurance purposes during the crisis.

Chart 5: The market for central bank reserves under the current ‘floor’ framework

![Diagram showing market for central bank reserves under the current ‘floor’ framework.]

Chart 6: Gap between Bank Rate and market overnight interest rates

![Diagram showing gap between Bank Rate and market overnight interest rates.]

Sources: Bloomberg and Bank calculations

11 Market rates are slightly below Bank rate because most overnight deposits with SMF participants come from businesses without access to reserves: see https://www.bankofengland.co.uk/quarterly-bulletin/2018/2018-q1/sterling-money-markets-beneath-the-surface.
Waiting for the exit: (a) The MPC’s QT strategy

How will the balance sheet evolve in the future? That depends heavily on the MPC’s monetary policy decisions.

The MPC has been clear that its primary policy instrument is Bank Rate, which can move in either direction from its current level of 0.75%. If the MPC chooses to ease policy and expand the balance sheet further as part of that, the Bank is operationally ready to do so.

But at some point, as part of a future tightening strategy, the time will come to start reducing the stock of purchased assets. Given the sums involved, this is not something to do lightly – so the MPC has set out the steps it will follow:

- First, the MPC does not intend to begin QT until Bank Rate has risen to a level from which it could be cut materially if required. The MPC currently judges that to be around 1.5%.\(^{12}\)

- Second, QT will be conducted over a number of years at a gradual and predictable pace, chosen by the MPC in light of economic and financial market conditions at the time.

- Third, the QT path will take account of the need to maintain the orderly functioning of the gilt and corporate bond markets including through liaison with the Debt Management Office.

- And, fourth, the QT path can be amended or reversed as required to achieve the inflation target.

When might this all start? No time soon, if you ask the financial markets! The current forward yield curve does not reach 1.5% at all (Chart 7). But options markets price in a small probability of it occurring, and (as the chart shows) expectations can shift quite rapidly: less than a year ago the implied central case start date was in 2021.

\(^{12}\) This judgment was adjusted down from around 2% in June 2018, reflecting revised estimates of the effective lower bound for Bank Rate. See [https://www.bankofengland.co.uk/-/media/boe/files/monetary-policy-summary-and-minutes/2018/june-2018.pdf](https://www.bankofengland.co.uk/-/media/boe/files/monetary-policy-summary-and-minutes/2018/june-2018.pdf)
Waiting for the exit: (b) Delivering our policy goals in the early phase of QT

Our job on the Bank executive is not to decide when to start QT, but to be ready to execute it when asked to do so, and in a way that ensures that the Bank balance sheet continues to deliver all of our objectives. Specifically, that means that we must be able to: maintain control of short-term interest rates; keep the balance sheet flexible enough to meet the financial stability needs of the system (whether through liquidity insurance or opening up the balance sheet to new participants as the financial sector evolves); and remain ready to implement unconventional monetary expansion if the economy were to deteriorate.

Just as QE increased the quantity of central bank reserves, QT will reduce it. In the early stages of QT, whilst reserves are still plentiful, the Bank’s policy goals are likely to be best met by maintaining the current floor system. But a floor only works when there is a sufficient supply of reserves. At some point, QT will drive reserves below the level reserves holders are comfortable holding at the prevailing policy rate, and that scarcity will begin to push short term market rates up above Bank Rate – and make them more volatile – as firms bid up to borrow between themselves. At this point, if we were to do nothing, the existing framework would cease to maintain interest rate control (Chart 8).13

13 Market rates would not rise without limit: under the Bank’s ‘Operational Standing Facilities’, firms can borrow against high quality collateral at a spread of 25 basis points above Bank Rate.
Waiting for the exit: (c) Ensuring a smooth transition. Findings from the Discussion Paper

To judge where this instability might kick in, we need to understand why firms demand reserves – what they use them for and how readily they would swap them for other liquid assets in different states of the world. In economists’ language, we need to understand the shape of the demand curve for reserves.

The difficulty is that this concept is not directly observable. That’s why we issued our Discussion Paper, as a way of beginning a conversation with firms. Since then we have met a broad and representative range of our counterparties: large and small, simple and complex, old and new, UK-based and overseas businesses. All told, we spoke to firms accounting for about three-quarters of current reserves balances.

What did we find? Unsurprisingly, different firms thought about their reserves holdings in quite different ways – and no-one used the term ‘demand curve’! Nevertheless, we can build up a picture of demand for reserves by understanding three core drivers of demand, common to all firms.

First, at the most basic level, **firms hold reserves to meet both expected and unexpected sterling payments during each day**. The shape of this demand depends on firms’ business models. Calendar effects are important for retail banks, for example; whereas wholesale banks and gilt-edged market makers are more affected by the ebb and flow of capital markets activity. And access to (and efficiency of) payment and settlement systems can have important implications for the reserves that firms need to hold. So innovations such as the Bank’s RTGS renewal programme\(^\text{14}\) have the potential to transform demand levels.

Second, **reserves are an important form of insurance against liquidity outflows over short, but multi-day, horizons**. Regulatory rules require firms to hold liquid assets to cover projected outflows over a

---

\(^{14}\) [https://www.bankofengland.co.uk/payment-and-settlement/rtgs-renewal-programme](https://www.bankofengland.co.uk/payment-and-settlement/rtgs-renewal-programme)
thirty day period. This doesn’t have to mean reserves – firms could instead hold liquid government bonds, mortgage backed securities or even equities. But, unlike reserves, such assets have to be liquidated in the market before they can be used to meet outflows. Of the firms we spoke to, even those with strong market access showed a marked reluctance to rely on that access as their only source of liquidity, particularly to meet projected needs in the initial days of a liquidity stress. For that, they preferred to hold reserves.

Taken together, these two factors make up what I will call a ‘preferred minimum range of reserves (PMRR)’ for firms. Chart 9 illustrates the aggregate PMRR for the market as a whole, defined as the region where reserves begin to become scarce. We expect market interest rates to be relatively stable around Bank Rate in this region. This range will naturally include any buffers that firms are holding above their true minimum needs to deal with uncertainties over potential stressed outflows, access to wholesale markets and so on.

The third driver of demand is the rate of return firms can earn on reserves relative to other liquid assets. In the current environment of very low bond yields, for example, reserves look quite attractive relative to gilts and other liquid assets, encouraging firms to hold reserves well in excess of their minimum preferred levels.15 But QT, when it comes, may imply a steeper yield curve than we see today. And that would incentivise firms to term out their liquid asset holdings towards gilts and other securities and reduce their reserve levels back towards their individual PMRRs.

An important question is whether this increased relative attractiveness of alternative liquid assets would also reduce the PMRR. Few of the firms we spoke to expected a big effect on their minimum holdings. That’s consistent with more detailed surveys carried out by the Federal Reserve, which suggested little response.

---

15In practice, of course, asset prices are jointly determined with reserves demand: they will have adjusted to ensure that the aggregate supply of reserves, determined by asset purchases, is willingly held.
even to quite large changes in the opportunity cost of holding US dollar reserves. But we should be cautious of these findings when alternative rates of return have been so low for so long.

**Bringing all this together, our conversations with firms suggest the current sterling PMRR is of the order of £150-250bn.** Table 1 breaks this down into a base estimate of firms’ ‘true minimum’ demand to meet payments and liquidity outflow needs, and an additional estimate of the buffers held above these minima. These estimates draw on firms’ descriptions of their preferred holdings and their management of fluctuations in reserves levels, cross-referenced with our own analysis of historic reserves account activity.

**Table 1: Current estimate of the UK PMRR based on feedback to the Discussion Paper**

<table>
<thead>
<tr>
<th></th>
<th>Estimate (£bn)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PMRR base</td>
<td>120 – 160</td>
</tr>
<tr>
<td>PMRR buffer</td>
<td>30 – 90</td>
</tr>
<tr>
<td><strong>TOTAL PMRR</strong></td>
<td><strong>150 - 250</strong></td>
</tr>
</tbody>
</table>

I should stress that there is no mechanical link between this figure and the quantity of gilts that the MPC will need to sell during QT: it does not for example mean that gilt sales must stop when reserves levels hit the PMRR. As I will come on to describe, the whole point of the balance sheet framework that we are proposing is that it will allow the size of the balance sheet to be determined independently from the quantity and composition of QE and QT.

Adding back the current level of banknotes and other non-reserves liabilities to the PMRR, and assuming nothing else changes, suggests that a steady-state balance sheet after QT of £275-375bn (12-18% of GDP) would be needed to deliver the Bank’s monetary policy and financial stability policy objectives. That’s materially higher than pre-crisis levels. But it’s roughly half where we are today (Chart 10), and narrower than we’ve previously indicated, reflecting the more detailed information available from the Discussion Paper feedback. As one cross-check, the required reduction in reserves to get to the PMRR is proportionally similar to that estimated for the United States by the Federal Reserve.

---


17 No attempt is made to estimate any potential future rate of return effect.


19 An April 2019 speech by Lorie Logan, Senior Vice President of the New York Federal Reserve cites an estimate for the US ‘Lowest Comfortable Level of Reserves’ of $800-900bn, which is around 30% of peak reserves ($2.8tn). By comparison, our PMRR estimate is 30-50% of current UK reserves. See https://www.newyorkfed.org/newsevents/speeches/2019/log190417 for details of the US estimate.
Chart 10: Estimated range for steady-state post-QT balance sheet

At the same time, it is worth underscoring how uncertain this estimate of the PMRR still is. Like all informal survey-based estimates, it is subject to sampling error. It reflects a varied set of responses from firms – many of whom are themselves, by their own admission, genuinely unsure how to calibrate their own future demand after such a long period of abundant supply and low interest rates. And the PMRR will certainly adjust over time in response to changes in external conditions. Factors which might reduce the future PMRR include higher rates of return on alternative assets, or greater access to and confidence in the functioning of short-term money markets. Factors which might increase the future PMRR include greater risk aversion or further expansion in the range of firms with access to reserves accounts. For example the Bank recently announced that it would be consulting on the appropriate level of access for new payments providers.20

Dealing with uncertainty: post-QE frameworks

We can, and will, continue to strengthen our analysis of the demand for reserves, through more structured and frequent dialogue. But if there’s one big message I take from our Discussion Paper exercise, it’s that, no matter how deep we dig, we will still face persistent and material uncertainty about the future demand for reserves. We must ensure our post-QT framework can deal with that uncertainty robustly if we are to be confident that we can continue to use our balance sheet to deliver monetary and financial stability.

One response would be to stick with the simple floor, but stop contracting reserves well before we hit the PMRR, leaving an ample buffer to ensure market interest rates are stable (Chart 9). This is the approach

adopted by the US Federal Open Markets Committee (FOMC). The assets of the Federal Reserve have fallen by over $600bn since US QT began in autumn 2017 – and US reserves are now more than 40% below their 2014 peak (Chart 11).

Chart 11: QE and QT in the United States

We know from our own experience that this ample reserves floor system can be an effective way to deliver unconventional monetary policy whilst also maintaining control of interest rates. It is particularly suited to the US context, where the supply of reserves is subject to relatively large day-to-day variation as government deposits with the central bank and other items fluctuate. But a key implication of this approach is the need to maintain an estimate of when reserves will become scarce, in order to keep reserves supply well above that level. As we have seen, this estimate is intrinsically uncertain.

The market impact of the early stages of US QT was extremely smooth, assisted by clear communications in advance about the way in which the Federal Reserve’s assets would run off progressively over time. But risky asset prices began moving more sharply in the later part of 2018 and into 2019, and there was persistent upward pressure on some short-term money market rates relative to official rates (Chart 12). Successive FOMC minutes recorded debates over how far these developments might reflect the impact of falling reserves levels and balance sheet normalisation more broadly.

---

For the FOMC’s current Balance Sheet Normalization Principles and Plans, see https://www.federalreserve.gov/monetarypolicy/policy-normalization.htm

The larger and more volatile gap between total assets and the central bank reserves funding them in the US reflects a more important role for non-reserves liabilities, including banknotes (which account for over 40% of the balance sheet, reflecting global demand) and government deposits.
Published Federal Reserve analysis and policymakers’ discussion\(^{23}\) suggested that reserves remained clearly in the ‘ample’ zone, and ascribed upwards pressure on short-term interest rates to unrelated factors, including the changed pace of Treasury bill issuance and changing dynamics in the repo market.

Nevertheless, in light of broader economic and financial conditions, the FOMC decided to announce a somewhat earlier end-date for normalisation than market participants had been expecting at the start of QT. That is expected to leave reserves initially ‘somewhat above the level … necessary to efficiently and effectively implement monetary policy’. Reserves will then be allowed to decline gradually to that level over time, as currency and other non-reserve liabilities grow.

\begin{chart}
\textbf{Chart 12: US short-term interest rates since 2014}
\end{chart}

\begin{itemize}
\item Source: Bloomberg and Bank calculations
\end{itemize}

In our Discussion Paper, we described a somewhat different approach to handling the uncertainty intrinsic to estimating reserves demand. At an appropriate moment during the QT process, we proposed to switch from using a simple floor system to a framework in which we allow reserves holders to determine how many reserves they need. Under this approach, the Bank would stand ready to lend reserves, at Bank Rate, in unlimited amounts against high quality collateral, through periodic Open Market Operations (OMOs) during the QT phase. Such an approach does not seek to ‘oversupply’ reserves in order to maintain a floor system, and so avoids the need to estimate the point of reserves scarcity.

The Bank’s balance sheet should naturally be somewhat smaller under this system than it would be under an ‘ample reserves’ floor. That’s partly because we would not need to build in a buffer to account for uncertainty over the true level of the PMRR. But it’s also because the balance sheet cost for banks obtaining reserves via OMOs is likely to require there to be a degree of reserve scarcity before banks will participate.\(^{24}\) As a result, the quantity of reserves should settle within the PMRR, where some reserves scarcity is evident.


\(^{24}\) Because they are secured funding transactions, OMOs will involve a leverage exposure, and will be subject to haircuts, which lower the value of reserves lent relative to the collateral supplied.
A benefit of our approach is that it allows choices over the asset composition of our balance sheet (eg between OMOs and outright holdings) during QT. At the same time, we retain one of the key benefits of a floor system, which is the capacity to allow for the smooth expansion in assets and reserves in the future, whether for liquidity insurance purposes or for monetary policy goals as directed by the MPC, without interfering with interest rate setting.

At least two design challenges remain, however.

The first is that we cannot perfectly anticipate how short-term market interest rates will behave in the new system, and will therefore need to be responsive in our operational approach:

- It is likely that market rates will be a little higher on average relative to Bank Rate than they are now, reflecting the fact that reserves will be in somewhat scarce, rather than ample, supply. As a one-off shift, this is unlikely to be of great significance to the setting of monetary policy.

- Market rates may be slightly more variable too, particularly initially as banks adjust to the new system and learn to identify and bid for the amount of reserves they need.

Given the importance of achieving control of interest rates, it will be important to be ready to vary the nature of our operations if either effect becomes problematic, for example by: running more frequent OMOs; varying the terms of those OMOs; or amending our standing facilities. We take some comfort from our discussions with banks that their demand for reserves within our framework should keep us on the relatively flat part of the demand for reserves curve, limiting overnight rate volatility. Nevertheless we will continue to deepen our understanding of the demand curve and its drivers through our own analysis, and through further engagement with the market.

A second challenge is to ensure that the new system, when in place, works for all market participants. The range and number of participants may well increase over time, and market structures may need to evolve to support the effective transfer of reserves through the system. So we will prioritise work on OMO design to ensure that SMF participants of all types, sizes and sophistication have ready access to reserves. In doing so, we will need to take care not to impede developments in repo and other markets unduly.

Conclusions

The Bank’s balance sheet will not be returning to its pre-crisis form, even after QT. That reflects the role it needs to play in delivering the Bank’s monetary policy and financial stability objectives in a new environment. But the balance sheet will likely shrink materially as the APF is unwound, and our best current estimate of the level of reserves under our proposed framework is in the range £150-250bn. As I mentioned earlier, this
number gives no indication about the extent of QE sales, because our proposed framework allows that to be
determined independently from the question of our medium term balance sheet size.

The Discussion Paper we issued in August last year was intended to stimulate debate – and from that
debate improve our understanding of how the Bank of England’s balance sheet should be managed in the
future.

It succeeded in doing so. Conversations with our counterparties have reinforced our intention to employ the
framework for controlling short term interest rates that we proposed in 2018. But it also helpfully raised
questions about the challenges we and participants will face during this period.

As a result we will be engaging carefully on the factors that will need to be taken into account in the design of
our future open market operations.

I hope my comments today will encourage thinking on this topic more broadly.

At times it can appear from the history of central banking that, when it comes to implementing policy,
practice leads theory by a long way. But practice always has to be underpinned by a clear grasp of reality:
this is the only way to achieve the Bank’s monetary and financial stability objectives. And that is how we will
continue to manage the Bank’s balance sheet in the years to come.

Thank you