Box B   The long and uncertain road to exiting Quantitative Easing

By Cyrille Lenoël[1]

There is a rich literature on Quantitative Easing (QE) but less so on its unwinding.[2] QE was initiated in 2001 in Japan, followed by the US in 2008 after the Global Financial Crisis, and there has been time to evaluate it, while there has not been any sustained unwinding of QE so far[3]. The only episode that approaches an unwinding of QE is the so-called 'Taper Tantrum' in the United States in 2013, when Treasury yields surged on the news that the Federal Reserve would be slowing down its purchases of bonds.

In this box, we survey the literature on QE and use the findings to discuss possible exit strategies. While there is little consensus among central bankers on the optimal exit strategy, what stands out is that the road to unwinding will probably be a long and uncertain one, and central banks may even keep large amount of government bonds permanently on their balance sheet.

The channels of QE

Quantitative easing is the process whereby a Central Bank (CB) purchases government or corporate bonds to stimulate the economy. Purchases are done in the secondary market in order to prevent 'monetary financing', which would be when the CB directly buys bonds issued by the Treasury (Macchiarelli and McMahon, 2020). The purchases tend to reduce the yields on the bonds and thus are intended to feed through to lower interest rates for households and businesses, stimulating economic activity in a similar way as a cut in the policy rate. In the UK, the Bank of England has announced the purchase of £895 billion worth of bonds between November 2009 and November 2020.

QE effects are generally decomposed in the economic literature into three channels: signalling, portfolio rebalancing and liquidity premium (Bailey et al., 2020). When the CB announces its intention to do QE, it also announces a timetable i.e. the total value of bonds to be purchased and over what period – usually several years. Because long rates are closely related to the expected path of short rates, such an announcement is seen by markets as a commitment to ease monetary policy for a significant period, and yields start declining at the announcement date, rather than when the CB actively starts purchasing bonds. This is the **signalling** channel.

On the other side of the bond transactions are sellers like money market funds or pension funds. These will generally reinvest the proceeds from their sale into other assets with higher yields like shares or properties. That process will in turn reduce the yields of other asset classes, making the households that hold these assets wealthier and able to spend more. This is the **portfolio rebalancing** channel. The academic literature models this channel by replacing the Efficient Market Hypothesis with assumptions of portfolio preferences so that different assets are imperfect substitutes on account of their non-pecuniary properties (Brainard and Tobin (1963), Andres et al. (2004), Chen et al. (2012) and Harrison (2012)), portfolio adjustment costs (Harrison, 2011, 2017), or preferred habitats, in which investors might demand certain assets for specific – perhaps regulatory – purposes (Vayanos and Vila, 2009, 2020).

The third channel is the liquidity premium. A bond investor will occasionally need to sell some of the bonds it holds and the risk that it may not find a willing buyer on time is called the **liquidity premium**. This liquidity premium is incorporated in the price of bonds. The fact that the CB becomes a willing buyer of a large quantities of bonds reduces the liquidity premium and therefore yields. The liquidity channel relies on the existence of a market or informational friction, which creates a role for central bank asset purchases in encouraging trading and reducing liquidity premia in a given market (Joyce et al., 2011; Haldane et al., 2016). By meeting the increased demand for safe and liquid assets by the banking sector and acting as a substitute for private sector collateral QE can also support the bank lending channel (Corrado et al., 2020).[4]

Diminishing returns?

A difficult question for a CB initiating QE is how to evaluate the amount of bonds necessary to purchase in order to reduce yields to the target level. Is it possible that QE may suffer from diminishing returns whereby the central bank has to buy ever increasing quantities of bonds to have the same marginal effect.

Some studies have found diminishing effects of QE in later rounds – see for example Greenlaw et al. (2018) and Krishnamurthy and Vissing-Jorgensen (2011). One argument for diminishing returns is that the term premium – defined as the difference between the bond yield and the average expected value of future short-term interest rates over the life of the bond – may not decline too much without distorting markets. Via the portfolio rebalancing channel, QE reduces the term premium, which may then become negative. For example, with German Bund yields currently being negative, investors have the option of holding paper currency instead of buying Bunds, which may limit the effect of QE if the European Central Bank wants to increase its QE programme. However, NIESR's term premium estimate for Germany has been negative since May 2019, and it does not seem to have produced market dysfunctions in Germany or the Euro Area so far (National Institute Term Premium Tracker, June 2021)

On the other hand, additional QE bond purchases may remove bonds from investors who are more reluctant to sell them and thus who demand ever higher prices (and lower yields), suggesting the possibility of non-diminishing returns. In an empirical study, Ihrig et al. (2018) find persistent effect for all rounds of QE in the US.

There are differing views from policy makers about when QE is most effective. Bailey et al. (2020) and Vlieghe (2021) from the Bank of England have argued that QE is particularly effective in crisis times. As expressed by Gertjan Vlieghe "[QE] is a very powerful tool to lower yields when market functioning is poor, by significantly increasing aggregate liquidity through abundant reserves and signalling the willingness to offset shocks. But when market functioning is restored, and if long term yields already at very low levels with inflation expectations near the target, in my view the ability for QE to impart additional macro-economic stimulus is limited. In other words, the impact of QE is state-contingent" (Vlieghe, 2021). But Ben Bernanke, former Federal Reserve chair, argues that "the research rejects the notion that QE is only effective during periods of financial disruption. Instead, once market participants' expectations are accounted for, the impact of new purchase programs seems to have been more or less constant over time, independent of market functioning, the level of rates, or the size of the central bank balance sheet" (Bernanke, 2020).

Quantitative estimates

Figure B1 shows the minimum, maximum and median estimates from the literature of QE on 10-year yield reduction, normalized to purchases of 10 per cent of GDP, for the US, UK, Japan and Euro Area. The vast majority of the studies are for the US, but there are three for the UK and Euro Area each, and two for Japan. The studies unanimously conclude that QE lowers bond yields significantly, but the range of the estimates is quite wide. In the case of the UK, studies find that a QE expansion of 10 per cent of GDP reduces 10-year gilt yields by 46.5 basis points. Applying this estimate to the £895 billion of announced bond purchases by the Bank of England between 2009 and 2020 (or 47.5 per cent of average GDP), suggests a cumulative decline in 10-year gilt yield of 2.2 percentage points. Over the same period, the yield declined from 3.8 per cent to 0.3 per cent, which suggests that 2.2 percentage points of the 3.5 percentage points decline can be attributed to the QE programme.[5]

Figure B.1 Literature estimates of effects of QE bond purchases on 10-year yields
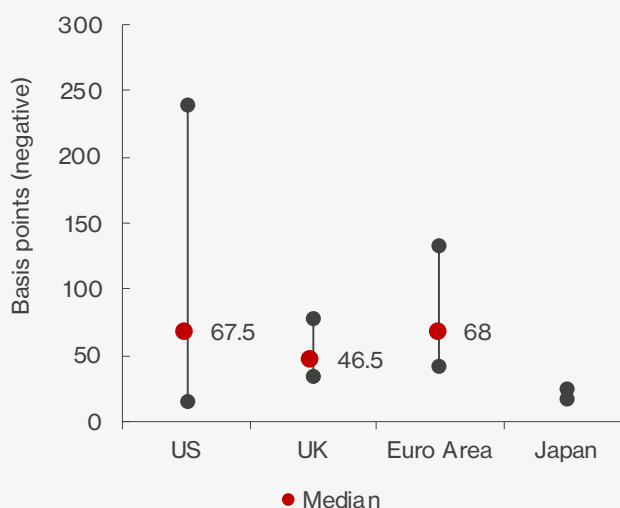
Note: Gagnon (2016) and author's calculations. Purchases are normalized to 10 percent of GDP.

---

5   The empirical evidence on the effects of QE on low frequency macroeconomic variables like inflation and output is less conclusive. See Rossi (2021) for an excellent review of this literature.

Exiting QE

There are mainly two arguments for unwinding QE. The first one is to 'normalise' monetary policy so that there is more room to ease when the next negative shock hits the economy (Chadha, 2017). The second is that the recent rise in inflation in the UK may become more persistent if accompanied by a strong demand-side recovery and a rise in wages, and the Bank of England should consider tightening the overall policy stance as soon as next year. The two arguments are, of course, not mutually exclusive.

By reversing the channels of QE discussed above, one can get an idea of the likely effects of unwinding QE. The Taper Tantrum episode provides support for the idea that unwinding QE should be done over several years and in moderate steps, to prevent a strong market reaction. The greater uncertainty about the efficacy and channels of QE compared to the well-known effects of short-term policy rates, also argues for gradualism in unwinding it (Williams, 2013). An announcement to slow down asset purchases and start reducing the stock of holdings would be a strong signal to markets that the CB will tighten monetary policy and not look through a rise in inflation. Gradually reducing bond holdings will reduce liquidity in the government bond markets, and financial intermediaries need to be prepared to see one of the largest participants in this market step back. The portfolio rebalancing channel will lead to higher bond yields and term premia.

The interaction between policy instruments

The quantitative effects of entering and exiting QE may not be symmetrical because of the interaction with policy rates. QE was set up as a complementary easing instrument when policy rates were believed to be at the Zero Lower Bound (ZLB). But now that central banks have two main policy tools (policy rates and QE), it is not clear which should be used first when tightening monetary policy. In the case of the UK, if we assume the same median estimate of QE (46.5 basis points), then the hypothetical case of a full reversal of QE by the Bank of England could increase the 10-year gilt from a current yield of 0.5 per cent to 2.7 per cent. However, there is considerable uncertainty around this estimate not only because of the uncertainty regarding how much the size of the balance sheet will be reduced but also because of the interaction with the policy rate. For example, if Bank Rate is raised first so that it is not at the ZLB, the impact of unwinding QE on yields may be lower per se.[6]

In its June 2018 meeting, the MPC set out its policy for unwinding QE: the Bank's balance sheet should be unwound "over a number of years at a gradual and predictable pace", allowing reserves to fall back to a level demanded by commercial banks as evidenced through participation in regular repo operations. The MPC also declared its preference for increasing Bank Rate first to 1½ per cent, before beginning to reduce its balance sheet. The threshold of 1.5 per cent, while somewhat arbitrary, was viewed as a level from which Bank Rate could be cut materially (or raised further) as necessary. This approach allows Bank Rate to be used as the primary instrument to set the stance of monetary policy in response to shocks in either direction, while a gradual and orderly balance sheet unwind continues. Broadbent (2018) justified this approach by reference to the fact that Bank Rate is a more flexible instrument, which can be adjusted more nimbly to shorter-term macroeconomic shocks, with more predictable effects.

While this policy has the merit of providing clear guidance, it has not been tested against alternative policies of normalisation: either to reduce the balance sheet before increasing Bank Rate, or to act simultaneously on the balance sheet and rates.

There are arguments for reducing the balance sheet first. QE may be more distorting to financial markets than the standard policy rate. By affecting both the short end of the yield curve (with the policy rate) and the long end (with QE), the current policy stance tries to some extent to control the yield curve, which affects an important price signal for financial markets (Chadha, 2021). QE also has an impact on the profitability of some financial intermediaries like money market funds and banks, which may create financial instability and asset price bubbles. Darracq-Paries and Kuehl (2017) explain that frictions in financial markets make QE particularly effective at easing monetary policy at the ZLB via the term premium, and the corollary is that it is optimal to unwind QE before increasing policy rates in order to reduce the welfare costs of portfolio frictions.

---

6   Another complication regarding the exit from QE and a rise in interest rates relates to its possible impact on public finances given that the share of government debt held by the Bank of England is expected to reach some 40 per cent. The Treasury has received so far an indemnity of £112 billion from the Bank of England from marked-to-market gains associated to the gilts it holds, but is at risk of having to compensate the Bank if yields increase. See Macqueen (2021) and Allen (2021) for a detailed discussion.

National Institute of Economic and Social Research

The recent rise in inflationary pressures requires central banks to start preparing the ground for a normalisation in monetary policy. Given the role of supply-side factors in pushing inflation up and continued uncertainties regarding the pandemic, communication around tapering asset purchases and policy rate normalisation will be crucial to avoid a significant tightening in financial conditions which risks the ongoing recovery.

REFERENCES

Allen, W. (2021) 'Managing the fiscal risk of higher interest rates', NIESR Policy Paper 25, March

Allen W.A., Chadha J.S. and Turner P. (2021) "Quantitative Tightening: Protecting Monetary Policy from Fiscal Encroachment" National Institute of Economic and Social Research Policy Paper 027

Andrés, J, López-Salido, J and Nelson, E (2004), "Tobin's Imperfect Asset Substitution in Optimizing General Equilibrium", Journal of Money, Credit and Banking, 36(4), pages 665–90.

Bailey, A. J., Bridges, J., Harrison, R., Jones, J., & Mankodi, A. (2020). The central bank balance sheet as a policy tool: Past, present and future. Bank of England Working Paper No. 899

Bank of England (2018): The Bank of England's future balance sheet and framework for controlling interest rates. Discussion Paper August 2018

Bernanke, B. S. (2020). The New Tools of Monetary Policy: American Economic Association Presidential Address. Brookings Institution, January, 4.

Blinder, A. S. (2010). Quantitative easing: entrance and exit strategies. Federal Reserve Bank of St. Louis Review, 92(6), 465-479.

Brainard, W. C., & Tobin, J. (1963). Financial intermediaries and the effectiveness of monetary controls. American Economic Review, 53(2), 383-400.

Broadbent, B (2018), "The history and future of QE", speech given at the Society of Professional Economists, London. 23 July

Butt, N., Churm, R., McMahon, M. F., Morotz, A., & Schanz, J. F. (2014). QE and the bank lending channel in the United Kingdom. Bank of England Working Paper No. 511

Chadha, J. S (2017), Commentary: Interest rate normalisation, National Institute Economic Review No. 241, August 2017.

Chadha, J.S. (2021), Written evidence submitted to the Quantitative Easing Inquiry by Economic Affairs Committee https://committees.parliament.uk/writtenevidence/36897/html/

Chadha, J.S., & Holly, S. (2011). Interest rates, prices and liquidity. Cambridge University Press.

Chen, H, Cúrdia, V and Ferrero, A (2012), "The macroeconomic effects of large-scale asset purchase programmes", The Economic Journal, Vol. 122, issue 564, pages F289–F315.

Corrado, L., Schuler, T. and Chadha, J.S. (2020) Bank Reserves and Broad Money in the Global Financial Crisis: A Quantitative Evaluation. ECB Working Paper No. 20202463

Darracq-Paries, M and Kuehl, M (2017), The Optimal Conduct of Central Bank Asset Purchases, Deutsche Bundesbank Discussion Paper No. 22/2017.

Gagnon, J. (2016). Quantitative easing: An underappreciated success. PIIE Policy Brief, 16.

Giansante, S., Fatouh, M., & Ongena, S. (2019). Does quantitative easing boost bank lending to the real economy or cause other bank asset reallocation? The case of the UK. The Case of the UK (August 18, 2019). Swiss Finance Institute Research Paper, No.19-72.

Greenlaw, D, Hamilton,J, Harris, J and West, K (2018), "A Skeptical View of the Impact of the Fed's Balance Sheet", National Bureau of Economic Research Working Paper No. 24687.

Haldane, A, Roberts-Sklar, M, Wieladek, T and Young, C (2016), "QE: the story so far", Bank of England Staff Working Paper No. 624

Harrison, R (2011), "Asset purchase policies and portfolio balance effects: a DSGE analysis," chapter 5 in Chadha and Holly (2011)

Harrison, R (2012), "Asset purchase policy at the effective lower bound for interest rates", Bank of England Working Paper No. 444.

Harrison (2017) Bank of England Staff Working Paper No. 678 Optimal quantitative easing

Ihrig, J, Klee, E, Li, C, Wei, M and Kachovec, J (2018), "Expectations about the Federal Reserve's balance sheet and the term structure of interest rates", International Journal of Central Banking, Vol. 14(2), pages 341–90.

Joyce, M, Lasaosa, A, Stevens, I and Tong, M (2011), "The Financial Market Impact of Quantitative Easing in the United Kingdom", International Journal of Central Banking, Vol. 7(3): pages 113–61.

Kapoor, S., & Peia, O. (2021). The impact of quantitative easing on liquidity creation. Journal of Banking & Finance, 122, 105998.

Krishnamurthy, A and Vissing-Jorgensen, A (2011), "The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy", Brookings Papers on Economic Activity 42 (Fall), pages 215–65.

Kuang, C., Yang, J. and Zhu, W. (2020) Quantitative Easing and Bank Lending: The Liquidity Channel.

Macchiarelli, C. and McMahon, M. (2020) Quantitative easing and monetary financing: what's the difference?, Economics Observatory, 27 May 2020.

Macqueen, R. (2021) Interest rate rises and Covid-19 government debts, National Institute UK Economic Outlook, Spring 2021.

National Institute Term Premium Tracker, June 2021 available at https://www.niesr.ac.uk/sites/default/files/files/NIESR%20Quarterly%20Term%20Premium%20Tracker%20No_1%20-%20Full%20Document.pdf

Rossi, B. (2021) 'Identifying and estimating the effects of unconventional monetary policy: How to do it and what have we learned?', The Econometrics Journal 24:1

Vayanos, D and Vila, J (2009), "A Preferred-Habitat Model of the Term Structure of Interest Rates", National Bureau of Economic Research Working Paper No. 15487.

Vayanos, D and Vila, J (2020), "A preferred-habitat model of the term structure of interest rates", mimeo, London School of Economics.

Vlieghe, G. (2021) Speech at the University of Bath, on Thursday 27 May 2021

Williams, J (2013), "A Defense of Moderation in Monetary Policy", Journal of Macroeconomics No. 38, December, pages 137–50.