

DISCUSSION PAPER SERIES

DP12272

A CENTRAL BANK'S OPTIMAL BALANCE SHEET SIZE?

Charles A Goodhart

MONETARY ECONOMICS AND FLUCTUATIONS



A CENTRAL BANK'S OPTIMAL BALANCE SHEET SIZE?

Charles A Goodhart

Discussion Paper DP12272
Published 05 September 2017
Submitted 05 September 2017

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **MONETARY ECONOMICS AND FLUCTUATIONS**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Charles A Goodhart

A CENTRAL BANK'S OPTIMAL BALANCE SHEET SIZE?

Abstract

Unlike other facets of monetary policy renormalisation, there has been little discussion yet of what principles should determine the optimum size of a Central Bank's balance sheet, the end-point to which on-going portfolio reductions should approach. In this note I start by addressing the arguments of those who would leave this balance sheet very large, much as now; and then continue with the counter-arguments, also stressing the nature of the relationships between monetary and fiscal policies, and between the Central Bank and the Treasury's Debt Management Office.

JEL Classification: E50, E52, E63, H63

Keywords: Central Bank Balance Sheet, QE, Monetary Policy Renormalisation, liquidity, Debt Management, interest rate risk, auction risk

Charles A Goodhart - caegoodhart@aol.com

London School of Economics and CEPR

A Central Bank's Optimal Balance Sheet Size?

By C. Goodhart
Financial Markets Group
London School of Economics

Abstract

Unlike other facets of monetary policy renormalisation, there has been little discussion yet of what principles should determine the optimum size of a Central Bank's balance sheet, the end-point to which on-going portfolio reductions should approach. In this note I start by addressing the arguments of those who would leave this balance sheet very large, much as now; and then continue with the counter-arguments, also stressing the nature of the relationships between monetary and fiscal policies, and between the Central Bank and the Treasury's Debt Management Office.

JEL Classifications: E50, E52, E63, H63

Keywords: Central Bank Balance Sheet; QE; Monetary Policy Renormalisation; Liquidity; Debt Management; interest rate risk; auction risk.

Contact address:

Connaught House
London School of Economics
Houghton Street
London WC2A 2AE

Tel: 020 7955 7555

Email: caegoodhart@aol.com

Introduction

It is widely expected that the Fed will shortly announce that it will start to reduce the size of its balance sheet, by not reinvesting its inwards cash flows from interest payments and maturing principal. Such an announcement is predicted by commentators to be given at its September FOMC meeting. Other Central Banks, e.g. ECB and Bank of England, may follow on later. The sequencing of monetary policy normalisation, i.e. whether, and if so how far, interest rate increases should precede running down the balance sheet, has been widely discussed. And the speed of such balance sheet reduction has been, almost universally, proposed to be gradual, at least initially by not reinvesting cash inflows rather than by sales out of the Central Bank's portfolio. But what has been little discussed is what is the ultimate objective for the exercise; in equilibrium how large should the Central Bank's balance sheet be, and what are the principles that should determine that ultimate size? The purpose of this note is to address those questions.

One aspect of such a pull-back in portfolios is, however, generally agreed. Where the Quantitative Easing (QE) involved directional elements, to support credit flows through critical but weak markets, e.g. the mortgage market in the USA, such assets should eventually be entirely run off, leaving the direction of credit to market forces, and the assets left in the Central Bank's balance sheet should be entirely in the form of government debt. In so far as there are particular complications in this respect for the ECB, this note will not address them.

But what remains to be resolved is the ultimate size of its balance sheet that a Central Bank, having now inflated its balance sheet, via QE, by a multiple of four, or even five, times, should now aim for; indeed the criteria, the principles, on which such a judgment might be made seem unclear. What might determine the optimal size of a Central Bank's balance sheet? In this note I start by rehearsing the arguments of those who advocate keeping a very large balance sheet, perhaps roughly its present size. Then I shall discuss the counter-arguments of those who would like the balance sheet to revert to its previous lean level, representing the note issue and some minimal level of (perhaps unremunerated, but largely required) commercial bank reserves.

Arguments for Maintaining a Greatly Expanded Balance Sheet

The classic article in favour of a large balance sheet is by Greenwood, Hanson and Stein (2016), delivered at that August Jackson Hole meeting, now ably supported by a Citibank Research paper, 9 August 2017 (citivelocity.com), by Buitter, Rahbari, Jensen and Rojas. The arguments, briefly, are as follows. By paying interest on (excess) reserves (IOER) and/or offering reverse repos (RRPs), the Central Bank can continue to control the short-term official rate, and hence carry out its monetary policy mandate, whatever the size of commercial bank reserves and its own balance sheet. Meanwhile liquidity is desired, especially in a crisis, and having more liquidity enhances financial stability. But Central Banks can create such extra liquidity at a profit to the public sector, since their assets generally have a higher rate of return than their liabilities (seignorage), i.e. better than costless. So why not create much more? [Note that this line of argument can be extended to proposals for the Central Bank to extend electronic money balances and borrowing rights to a much wider set of clients, not only to non-bank financial intermediaries, but to everyone in the country. That raises broader issuers, which we shall duck here.]

It is true that a Central Bank's extension of its balance sheet tends to eliminate the need for, and the information from, other short term money markets, e.g. the interbank market. But the Central Bank still has that information on its books. As several economists have recently noted,¹ in a financial crisis, it is best *not* to have information on individual liquidity positions (of banks) publicly available, whereas the Central Bank, as Lender of Last Resort, does need such information, and has it more directly under a large balance sheet model. By the same token, the Calomiris/Kahn (1991) argument that tight liquidity enforces more prudent behaviour should (but would it?) be mitigated, or even reversed, if the Central Bank used its observation of sharply declining individual client balances to check what was happening and properly supervise micro behaviour.

GHS note that the role of QE in providing liquidity could be, more or less exactly, mimicked by the Treasury's Debt Management Office (DMO) running an equally massive 'Operation Twist',² issuing

¹ Notably Dang, Gorton, Holmstrom and Ordóñez (2017), and numerous other related publications by Gorton and Holmstrom.

² Operation Twist was an exercise tried in the USA in 1961, whereby large amounts of short-dated debt were sold, in order to hold up short-term interest rates, to protect the balance of payments, offset by large

vast quantities of short-dated Treasury Bills and retiring an equal amount of (much) longer dated debt. But they argue that it is preferable for the Central Bank, rather than the DMO, to do so, because this would eliminate ‘auction risk’, the risk that the auction of TBs might go badly.

Given this paean of praise for keeping the Central Bank’s balance sheet large, I had expected both GHS and BRJR to rehearse proposals for a further increase, perhaps 50% more, holding 30% of all government debt, instead of 20% as of now, (see FT, Wednesday, August 16, p. 1). Instead they end tamely, suggesting maintaining the current status quo. Thus GHS conclude, p. 387, that

“First, the Fed should keep a large balance sheet indefinitely going forward, even as rates rise well above the ZLB. While we do not attempt to pin down an exact dollar number, the current size of approximately \$4.5 trillion strikes us as a plausible baseline.”

And BRJR argue that ‘The optimal size of the balance sheet is unknown and probably unknowable’, pp 1 and 7. But while they present arguments for not shrinking a Central Bank’s balance sheet from its present level, they advance none for further increasing it. Since the present level was largely attained by happenstance, proposals to maintain that (ad hoc) level suggest a lack of determining principles, and, perhaps a certain doubt of the strength of the underlying arguments.

Arguments Against

In some large part because most of the analysis of QE, and of Central Banks’ balance sheets, has come from within the Central Banks themselves, the foremost concern expressed about their maintenance has been about the way that this could expose them to presentational and political pressures. Thus the likelihood that short-term interest would rise, as normalisation occurs, and so would lead to ever larger interest payments on their (sight) liabilities, primarily the reserve holdings of commercial banks, who are not politically beloved, while the interest payments on their longer dated assets would rise much more slowly. Depending on the details of the composition of their assets and liabilities, and on the accounting measures used (mark-to-market or hold to maturity), one could work out roughly what increases in interest rates might lead each Central Bank to post a

purchases of long-term debt, to hold down long-term rates, in support of the domestic economy. Whether this had much effect on the slope of the yield curve, or on the real economy, remains debatable.

loss, (a somewhat tedious exercise not attempted here). For rather obvious presentational reasons, Central Banks strongly dislike showing losses and will take accounting and practical steps to prevent that happening, (transparency is for others), see Goncharov, Ioannidou and Schmalz, (2017).

But it should be quite easy to take steps to prevent such a loss occurring at Central Banks, on account of maturity transformation, as interest rates rise back again, to normal or above. One simple way would be to reduce such maturity transformation by swapping their longer-dated, for shorter-dated, government debt. GHS suggest this, e.g. p. 363; W.A. Allen (2017) goes even further, suggesting that the longer-dated government debt (gilts) held in the UK's Asset Purchase Facility (APF) be all swapped for Treasury Bills (p. R67). Actually, it is not clear why this latter would make any real difference. All the profits, and losses, from the APF (net of running costs) already go to the Treasury. What is on the Bank of England's own balance sheet is loans to the APF to finance QE. Such loans either are, or could easily be made, at a margin above IOER, thereby completely protecting the Bank of England from such interest rate risk. Governor Mervyn King, who initiated QE1 in the UK and the Treasury's indemnity for the APF, was presciently aware of the accounting implications.

Another measure that would greatly reduce Central Bank, and Public Sector, exposure to such interest rate risk would be to pay a zero rate of interest on intra marginal reserves at the Central Bank, a step that would need to be accompanied by excluding such (intra-marginal) reserves from any leverage ratio calculation. Many Central Banks have changed the way that they remunerate commercial bank reserves, and several have either introduced, or at least considered, differentiating between the remuneration on marginal and intramarginal reserves, during the last decade. Further policy changes in this respect are possible; such present arrangements are not set in stone.

But, in an important, underlying sense, to focus on the effect on the Central Bank's Profit and Loss Account is to miss the point. Whether the accounting loss arising from rising interest rates, after the overall maturity of public sector has thus been shortened, is suffered in the Central Bank, an APF, or in the Treasury, it accrues to the public sector and to the taxpayers as a whole in any case. It is thus essentially a fiscal issue. GHS argue in response that the effect of QE (and thus the size of the Central Bank's balance sheet) on the *overall* duration of the government's debt could be maintained constant by having the DMO implement a 'barbell' strategy, whereby the short-dated liabilities

occasioned by QE are matched by massive issues of ultra-long debt. Pension funds and insurance companies would not be happy and would have to be bribed by higher rates to take up such a huge amount of ultra longs. In the one historical occasion that I can recall when government debt took up such a 'barbell' format, (the UK 1900-1914, when TBs and Consols formed the bulk of such debt), the debt ratio was much lower and conditions and context quite different.

What seems to me clear is that the overall duration and the acceptance of interest rate risk should be seen as relating to the public sector as a whole, not to separate segments of it. Having the Central Bank manage QE, and its vastly expanded Balance Sheet, according to one set of principles, and the DMO doing the rest, according to a different set of principles, is a recipe for conflicts of policy and principle. If the Central Bank is really going to maintain such a huge balance sheet indefinitely, then it should be given responsibility for all such debt management. In the UK this was the practice from 1694 until 1997, but always under the strategic guidance of the Treasury. But after the granting of Central Bank Independence, and, even more so, after the Bank of England was given responsibility for financial stability, financial supervision and macro- and micro-prudential measures, then giving back to it also powers to control the duration and interest-rate risk of the public sector debt is surely a step too far.

By the close of the period in which QE1 was in operation, by end 2009, the liquidity needs of the financial system in the UK and USA had been satiated. QE2 and QE3 were then purely fiscal, a reprise of Operation Twist. Whatever one might think about the relative success of QE2 and 3, (not much on this view), they should have been carried out by, and on the orders of, the Treasury's DMO, *not* by the Central Bank. If the Central Bank cannot persuade the Treasury to do this, it should not be done at all, *not* just on the Central Bank's sole initiative.

But what about the argument by GHS that the Central Bank, rather than the DMO, should undertake the expansionary arm of Operation Twist in order to eliminate 'auction risk', (pp 355-61)? In my view that is the weakest part of their paper. Given all the incentives to hold high quality liquid assets (HQLA), such as LCR, NSFA, etc., is there seriously much risk of an auction for short-dated (3 month and under) Treasury Bills failing? And to the extent that there was a perception of such risk, the private sector could be paid, or even forced, to underwrite such TB issues, as the Discount Houses

once did in the UK. Do we want to confuse the roles of the Central Bank and DMO to deal with such a (minor) issue as represented by 'auction risk'?

But QE1 was different, (and exercises to assess the effects of QE treating them all implicitly as similar are inappropriate). There was then a huge, and unsatisfied, demand for liquidity after the GFC, as measured, for example, by several liquidity premia, such as the margin between OIS and LIBOR. The need for liquidity was so pervasive that it required a build-up of bank reserves sufficient to drive overnight rates to zero to satisfy it fully.

Thus on liquidity management grounds, there is a strong case for the Central Bank to satiate the banking, and financial, system's demand for liquid assets, but not further. But what is the level of high quality liquid assets (HQLA) that would do this? Ricardo Reis (2016, p. 463) suggests that,

“Looking forward, keeping the market for reserves saturated is consistent with returning to a lean central-bank balance sheet. This means not zero excess reserves but rather closer to 1 trillion, or the size of the balance sheet in 2011, not pre-2008. The Federal Reserve can return to focusing on setting interest rates to control inflation, but now with its main target being the interest on reserves instead of the federal funds rate”

It is not clear how Reis arrives at his round number of \$1 tn. Perhaps the point is not so much the level, but the approach, i.e. that the Central Bank could, cautiously, proceed to reduce its balance sheet until signs of liquidity premia began to reappear. But when such liquidity premia might begin to resurface would depend not only on the remaining size of the Central Bank's balance sheet, but also on the availability of substitute HQLA, notably in the form of short-dated TBs. Consider a thought experiment; suppose that out of a balance sheet of 4X, 3X was swapped for TBs; what would change? This would be very different from having a Central Bank just run down 3X of assets without replacement by other liquid assets.

With excess reserves at the Central Bank no longer being counted against required leverage ratios in many regimes, and having no interest rate risk, it is arguable that 4X reserves is more liquid (and desirable) than 1X reserves and 3X TBs. But there must be some number Y, where 1X reserves, plus 3X + Y TBs, are just as liquid as 4X reserves. I would doubt if Y was a large number.

Given the continuing, and valid, concern about financial stability, there is, I believe, a fairly general consensus about the need to continue to satiate (the banking and financial system's) demand for liquidity. But that can be achieved by a whole variety of constellations of liquid assets. It does not all need to be done by massive holdings of commercial bank reserves at the Central Bank. It could, instead, be done with *much* reduced reserve balances offset by much larger holdings of short-dated TBs, (pace GHS' argument about 'auction risk').

Apart from auction risk, what are the arguments?

Does Liquidity Satiation Necessarily Imply Massive Excess Bank Reserves?

Under the Woodford/Lucas model, the transmission mechanism of monetary policy runs directly from the officially determined short-rate of interest, via expectations to longer rates, and thence to the real economy; the money supply is merely a memorandum item, and banks generally play no significant role. One does not have to be a full-blown Monetarist, however, (and I am not), to feel that a condition whereby a 100 b.p. cut in interest rates is accompanied by a 5% increase in broad money is likely to be much more expansionary than when the same cut in interest rates is matched by no increase in M3.

Financial intermediaries, especially banks, do, as their name suggests, also intermediate between the Central Bank and economic agents. Many agents primarily access the financial scene via such intermediation. If the transmission mechanism, via banks and other non-bank financial intermediaries, is seriously impaired, then so will be the efficacy of a given initial interest rate policy change, a point which most formal models miss.

One advantage of the old system, wherein reserve balances bore a zero interest rate, was that the desired R/D, reserve deposit ratio, was quite stable, keeping growth in base money in line with that of broad money. With IOER, especially when such reserves are excluded from required leverage

ratios, holding such reserves becomes a desirable asset class on its own, and the desired R/D ratio becomes highly variable. Hence the relationship between base money and the broad money stock breaks down, as evidenced in Table 1 below.

TABLE 1
THE COLLAPSE OF THE MONETARY BASE MULTIPLIER

	% change in		
	<i>H</i> monetary base	<i>M</i> broad money	Ratio <i>M/H</i>
USA			
2009	22.5	3.7	-15.3
2010	0.6	3.6	2.9
2011	31.2	9.7	-16.4
2012	2.9	8.2	5.1
2013	39.3	5.4	-24.3
2014	5.9	5.9	0.0
Japan			
2009	5.2	2.2	-2.8
2010	7.0	1.8	-4.9
2011	13.8	2.6	-9.8
2012	12.5	2.1	-9.2
2013	47.7	3.3	-30.1
2014	39.1	2.9	-26.1
UK			
2009	109.7	5.6	-49.6
2010	-1.3	5.5	6.9
2011	14.0	-3.1	-15.1
2012	50.5	0.2	-33.4
2013	7.6	0.7	-6.5
2014	1.5	-0.1	-1.6
Eurozone			
2009	-8.2	-0.5	8.3
2010	2.7	-0.7	-3.3
2011	25.1	2.2	-18.3
2012	22.7	3.0	-16.1
2013	-26.7	0.5	37.1
2014	-0.2	4.8	5.0

From Goodhart (2017).

It is true that, under the previous (corridor) system, with zero interest paid on reserves, in differing contexts a given cut in interest rates could also be accompanied by differing rates of growth of the monetary aggregates, both M0 and M3, but at least all the monetary aggregates would be telling the same story. One of the problems of the present system is that Central Bankers and commentators may assess the efficacy of their expansionary measures by looking at the scale of increase of the monetary base, the Central Bank's balance sheet, without realising that this may just have been absorbed into a liquidity trap, especially in the banking sector.

A problem with maintaining liquidity satiation is that the desired level of liquidity can alter dramatically and quickly between 'normal' times and the occasional crisis (panic). Providing sufficient liquidity (whether via reserves or TBs) *in normal times* in order to be confident of meeting panic-level requirements would probably require generally higher short-term interest rates to maintain price stability. What this suggests is that, in the longer term, rather than maintaining a hugely inflated amount of current HQLA, whether reserves or TBs, the authorities and banks should move towards a contingent, pre-positioned, capacity for banks to switch less-liquid assets into HQLA, with a hair-cut and interest rates, (high, but, of course, *not* penal), that would make such a switch unattractive during normal times, but positively attractive during panics, an idea proposed by Mervyn King (2016).

Seen in this latter light, the longer term objective should probably be to return to the status quo, pre 2008, with minimal, zero-interest-paying, bank reserves at the Central Bank, (plus a corridor system), supported by a much larger volume of HQLA, largely TBs, with a vastly increased, pre-positioned, arrangement for swapping less-liquid assets into TBs at the onset of any panic.

But one of the key remaining problems is the transitional flow problem of how to get from here, (the distorted present), to there (a better designed future).

Conclusions

1. Operation Twist (a.k.a. QE2 and QE3) should have been seen as primarily a fiscal issue. The effect on the long-term interest rate (good on the up; bad on the down) is counterbalanced by duration concern and interest rate risk, (bad on the up; good on the down). 'Auction risk' can be largely ignored or managed.

All Central Banks should follow the lead of the Bank of England in putting the bulk of their assets into an Asset Purchase (or Holding) Facility, financed by short-term loans from the Central Bank, with rates at a small margin over IOER. While the Central Bank should, *and will*, give its views on the implications for long-term rates, the final decision, on unwinding or increasing, should be taken by the Minister of Finance with the operations carried out by the DMO, *not* by the Central Bank. One minor advantage of switching responsibility to the DMO is that their communications tend to be lower-key and less dramatized and hyped than those of Central Banks.

My own view is that the disadvantages of the expanded balance sheet now outweigh the advantages, especially with long-rates remaining exceptionally low, so that the DMO should steadily run-down the APF at a gradual rate, to avoid 'tantrums' in the debt market, to zero. But whatever is decided, it should be done by the Minister, and DMO, after taking advice from the Central Bank.

2. A significant part of any remaining excess bank reserves, i.e. those not matched by Central Bank loans to the APF, should be immediately required to be switched into Treasury Bills. I would not expect this to have any significant effect on relative interest rates, but would be done both for presentational reasons and to explore the operational implications of running a system with much more HQLA in the form of TBs.
3. The size of Central Banks' balance sheets has become so large, and debt levels generally are so extended, implying great sensitivity to increasing interest rates, that it will take years, probably decades, to run off the excess reserves. During this extended inter-regnum, there is no alternative to running an IOER, floor system of rates.

But one should not lose sight, during this period, of the ideal end-point, which should include:

- a) Minimal, non-interest bearing bank reserves at the Central Bank, and a reversion to a 'corridor' system for short official rates.
- b) Much larger (than pre GFC) holdings by banks of HQLA, mostly in the form of TBs.
- c) A contingent, pre-positioning scheme for swapping banks' less-liquid assets into HQLA on terms carefully designed to be unattractive to banks during normal times, but attractive during panics/crisis.

References

Allen, W.A., (2017), 'Quantitative Easing and the Independence of the Bank of England', *National Institute Economic Review*, No. 241, August, pp 65-69.

Buiter, W., Rahbari, E., Jensen, K.D., and C. Rojas, (2017), 'Global Economic Views: Why does the Fed (or any central bank) want to shrink its balance sheet?', Citi Research, Global Economic Views, 9 August.

Calomiris, C.W., and C.M. Kahn, (1991), 'The Role of Demandable Debt in Structuring Optimal Banking Arrangements', *The American Economic Review*, Vol. 81, No. 3, June, pp 497-513.

Dang, T.V., Gorton, G., Holmström, B., and G. Ordonez, (2017), 'Banks as Secret Keepers', *American Economic Review*, 107(4), pp 1005-1029.

Financial Times, (2017), 'Decade of QE Leaves big central banks owning fifth of public debt', *Financial Times*, Wednesday 16 August, page 1.

Goncharov, I., Ioannidou, V., and M. Schmalz, (2017), '(Why) do central banks care about their profits?', June 2, available at SSRN: <https://ssrn.com/abstract=2979887> or <http://dx.doi.org/10.2139/ssrn.2979887>

Goodhart, C.A.E., (2017), 'The Determination of the Money Supply: Flexibility Versus Control', *The Manchester School*, Vol. 85, Issue 51, September, pp 33-56.

Greenwood, R., (2016), 'The Federal Reserve's Balance Sheet as a Financial-Stability Tool', Chapter in *Designing Resilient Monetary Policy Frameworks for the Future*, The Federal Reserve Bank of Kansas City, pp 335-397.

King, M.A., (2016), *The End of Alchemy: Money, Banking and the Future of the Global Economy*, London, UK: Little, Brown.

Reis, R., (2016), 'Funding Quantitative Easing to Target Inflation', Chapter in *Designing Resilient Monetary Policy Frameworks for the Future*, The Federal Reserve Bank of Kansas City, pp 423-478.