A monetarist view of the Fed’s balance sheet normalization period

Marcelo Veracierto

Introduction and summary

The Federal Reserve currently holds over $4 trillion in fixed-income assets. However, since the fall of 2017 the Fed has been in a “balance sheet normalization period,” during which the size of its balance sheet is gradually shrinking over time. In particular, the Federal Open Market Committee (FOMC) has been instructing the Federal Reserve trading desk to reduce its security holdings by reinvesting principal payments only to the extent that these payments exceed gradually raising caps. These caps on redemptions will be maintained until the Fed considers that its balance sheet has reached a desirable size.

At the same time, the Fed is implementing its target short-term interest rate by paying interest rates on reserves to depository institutions. It supplements this tool by offering overnight reverse repurchase agreement operations (ON RRP) to eligible financial institutions. By “encouraging competition, these instruments support interest rate control by setting a floor on rates, beneath which financial institutions with access to these facilities should be unwilling to lend funds” (Federal Reserve Bank of New York, 2018).

Given this monetary policy framework of administered interest rates, the Fed’s income statement may receive some pressure as interest rates increase during the normalization period. Moreover, as the size of the balance sheet decreases, the Fed’s interest income will shrink over time. A counterbalancing effect is that during the balance sheet normalization period, the stock of reserves will decrease as well, reducing the base over which interest is paid to depository institutions.

The Federal Reserve Bank of New York, which conducts open market operations on behalf of the FOMC, regularly provides projections for the System Open Market Account (SOMA) portfolio and net income. In all of its most recent reports (for example, Federal Reserve Bank of New York, 2016, 2017, 2018), the New York Fed has concluded that, under a variety of scenarios, the Fed’s remittances to the U.S. Department of the Treasury will remain positive during the balance sheet normalization period, avoiding the need to enter a deferred asset in its accounts. However, these calculations depend on how quickly reserves decrease over
time, which is closely tied to the rate of growth of currency in circulation. A key assumption that the New York Fed has regularly made in its previous projections is that currency will grow at the same rate as nominal gross domestic product (GDP). However, as interest rates increase, the demand for real cash balances will decrease over time. As a consequence, currency may grow at a slower pace than nominal GDP and reserves may decrease during the normalization period at a slower pace than previously estimated by the New York Fed. If reserves remain at higher levels than estimated, interest payments on reserves will be higher and the Fed’s net income will be lower than expected. In this article, I use estimates of money demand to redo the SOMA projections and evaluate whether remittances to the Treasury could be severely affected by a shrinking demand for real cash balances.

The first part of the article sketches the methodology previously used by the New York Fed and reviews the SOMA calculations. The second part of the article expands those projections by incorporating empirical evidence on the demand for money. However, the same basic result is still obtained: No red flags are raised in terms of Treasury remittances.

The last part of the article is concerned with the money multiplier, which is given by the ratio of some broad definition of money (such as M1, that is, the sum of currency in circulation, checking accounts, demand deposits, and negotiable order of withdrawal [NOW] accounts) to the monetary base (that is, the sum of currency in circulation and reserves). A potentially troubling result with the benchmark projections reported in this article is that they involve a striking doubling of the money multiplier over a period of three years. However, I argue that this should pose no difficulties since, as a first approximation, when the Fed pays interest on reserves, the size of its balance sheet becomes irrelevant for economic outcomes while the Fed gains full control of the money multiplier.

The Fed’s budget constraint

For simplicity I assume that the type of asset that the Fed owns is a bond of stochastic maturity that pays a coupon rate $k$ every period while maturity has not been reached. The per-period probability that the bond matures is $1 - \lambda$. The Fed’s budget constraint is then the following:

1) $\tau_t + i_t^R R_{t-1} + q_t (D_t - \lambda D_{t-1}) = \Delta R_t + \Delta C_t + k \lambda D_{t-1} + (1 - \lambda) D_{t-1},$

where $\tau_t$ are remittances to the Treasury, $C_t$ is currency in circulation, $R_t$ are reserves, $i_t^R$ is the interest rate on reserves, $D_t$ is the stochastic maturity bond, $q_t$ is the price of the bond, and $\Delta x_t = x_t - x_{t-1}$.

This stylized budget constraint is used to perform projections during the balance sheet normalization period. My first task is to review published SOMA projections. It happens to be the case that the latest New York Fed report that provides SOMA projections (Federal Reserve Bank of New York, 2018) does not provide enough detailed data for evaluating the extensions explored in this article (in particular, it does not report estimated SOMA interest expenses or interest income). An earlier report that provides such detailed data is Ferris, Kim, and Schlusche (2017). An apparent limitation of their analysis is that it considers a view of the balance sheet normalization period that is much simpler than the one that was later implemented. However, I view this as an advantage. The reason is that the introduction of gradually rising caps over which principal payments are reinvested complicates the analysis and is not central to the main focus of this article, which is how sensitive the SOMA projections are to incorporating an explicit demand for money. For this reason, I adopt the simpler scenario in Ferris, Kim, and Schlusche as my benchmark and refer to it as the “FKS scenario” for the rest of the article. I describe this scenario next.
The FKS scenario assumes that during the normalization period, the Fed pays interest on reserves and stops reinvesting in bonds when they mature. That is, it assumes that

2) \( i_t^R > 0 \), and

3) \( D_t - \lambda D_{t-1} = 0 \).

Thus, under the normalization period, the budget constraint of the Fed effectively becomes

4) \( \tau_t + i_t^R R_{t-1} = \Delta R_t + \Delta C_t + k\lambda D_0 + (1 - \lambda) \lambda^{t-1} D_0 \).

Under current practice, the Fed remits to the Treasury all of its net income on a period-by-period basis. That is,

5) \( \tau_t = k\lambda D_{t-1} - i_t^R R_{t-1} \).

From equation 4, it then follows that

6) \( \Delta R_t + \Delta C_t = D_t - D_{t-1} = -(1 - \lambda) D_{t-1} \),

that is, current contractions in bond holdings determine current contractions in the monetary base.

Once reserves reach $100 billion, the FKS scenario assumes that the Fed resumes asset purchases to support normal balance sheet growth. However, since the analysis that follows will focus on the period before that threshold would be reached, I use equation 4 throughout the rest of the article. Also, consistent with the FKS scenario, I assume the normalization period would have started in mid-2018.

FKS projections

Estimating the size of the balance sheet during the normalization period (as well as SOMA interest income) is far from trivial since it depends on the expected path of interest rates, which affects the pace at which agency holdings of mortgage-backed securities (MBS) pay down. Fortunately, FKS have already performed that calculation. The path for interest rates that they use is based on the Federal Reserve Bank of New York’s September 2016 Survey of Primary Dealers and is reported in figure 1. We see that starting from 0.8 percent in 2017, the fed funds rate is assumed to increase gradually, reaching 2.9 percent by 2021. The associated path for SOMA holdings that FKS estimate is reported in figure 2. We see that SOMA holdings decline slightly in 2018, when FKS assume that the balance sheet normalization period starts, and then quite abruptly through 2022 (after which the Fed resumes asset purchases to support normal balance sheet growth).

Given data for currency in circulation in 2016 and expected paths for inflation and GDP growth, a path for currency in circulation between 2017 and 2022 is then constructed assuming that

7) \( \frac{\Delta C_t}{C_t} = \frac{\Delta (PY_t)}{PY_t} \),

that is, that currency in circulation grows at the same rate as nominal GDP. The paths used for expected inflation and GDP growth rates are taken from the Federal Reserve Bank of New York’s September 2016
Survey of Primary Dealers and are reported in figure 3. Given the resulting path for $C_t$ (shown in figure 4) and for SOMA holdings $D_t$ (shown in figure 2), a path for reserves $R_t$ can then be obtained from equation 6. The path for reserve balances calculated by FKS is reported in figure 5. We see that reserves drop continuously over time, hitting the $100 billion threshold by 2022 (which triggers the resumption of asset purchases).

Figure 6 shows the path for SOMA interest expenses reported by FKS. We see that these interest expenses increase sharply, reaching a peak of $50$ billion before falling to $15$ billion by 2022. The basic reason why interest expenses remain relatively low even as interest rates increase is that reserves decline sharply over time (see figure 5). Given the SOMA interest income estimated by FKS and reported in figure 7, we obtain SOMA net income from equation 5. The resulting path is shown in figure 8. We see that SOMA net income decreases to $50.5$ billion by 2020 but then bounces back. As a consequence, remittances to the Treasury remain positive, and equation 5 remains valid (that is, the Fed is not forced to book a deferred asset).
Before proceeding to the next section, I need to obtain two important pieces of information from the FKS projections. Since in my simplified Federal Reserve budget constraint (equation 4) I impute all SOMA interest expenses to interest on reserves, I calculate the effective interest rate on reserves as follows:

\[ i_t^R = \frac{\text{SOMA interest expenses at date } t}{R_{t-1}} \]

where both denominator and numerator are obtained from the SOMA projections reported in figures 5 and 6, respectively. The resulting path for \( i_t^R \), which is reported in figure 9, is only slightly higher than the path for the federal funds rate in figure 1.
Another piece of information that will be extremely useful is the implied path for the monetary base, which is constructed as

\[ B_t = C_t + R_t, \]

where \( C_t \) and \( R_t \) are the currency in circulation and reserve balances reported in figures 4 and 5, respectively. The resulting path for the monetary base is reported in figure 10.

**A monetarist approach**

Given a fixed balance sheet path, any change in currency has to be offset by an equal and opposite signed change in reserves. This change in reserves in turn affects the total amount of interest payments on reserves and therefore the remittances to the Treasury. The previous section assumed that during the normalization period, currency grows at the same rate as nominal GDP. However, standard demand for money theory suggests that as interest rates increase, the demand for currency may grow at a much lower rate than nominal GDP or even shrink. As a consequence, reserves would not be able to contract as fast as previously calculated, and therefore interest expenses would be larger and remittances to the Treasury lower than the benchmark FKS calculations indicate. In order to address these concerns, in this section I redo the above calculations by imposing that currency growth must be consistent with a demand function for currency.

Lucas and Nicolini (2015) provide empirical evidence for the demand for currency during the 1915–2012 period. In particular, figure A1 in the appendix (which reproduces figure 2b from their paper) plots the inverse of the velocity of circulation of currency versus the three-month Treasury bill at an annual frequency. Given the stable demand function that this figure indicates, I postulate the following functional form for the demand for currency:

\[ \frac{C_t}{PY_t} = A_t \left(i_t + \phi_t \right)^{-\alpha}. \]
I then choose the parameters $A_c$, $\phi_c$, and $\alpha_c$ to fit the following three representative points of figure A1:

<table>
<thead>
<tr>
<th>$i$</th>
<th>$\frac{C}{PY}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.09</td>
</tr>
<tr>
<td>0.05</td>
<td>0.045</td>
</tr>
<tr>
<td>0.135</td>
<td>0.04</td>
</tr>
</tbody>
</table>

The required values for $A_c$, $\phi_c$, and $\alpha_c$ are 0.0316, 0.0002, and 0.1186, respectively, which imply an approximately constant interest rate elasticity of about $-12$ percent. Equipped with this demand function for currency, I turn to redoing the calculations of the previous section. The conditioning assumptions remain the same. In particular, the path for interest rates, SOMA holdings, SOMA interest income, and monetary base are left unchanged. The rationale for leaving SOMA holdings and SOMA interest income the same is that they are determined by the initial SOMA portfolio, the reinvestment decisions of the Fed during the normalization period, and the path for interest rates, and these are all unchanged across experiments. In turn, the path for the monetary base is left unchanged because, during the normalization period, it is strictly determined by the path of SOMA holdings and this remains the same.

With the increase in interest rates depicted in figure 1, figure 11 shows that the inverse of the velocity of circulation of currency given by equation 10 decreases quite significantly (instead of remaining constant as in the FKS experiment). As a consequence, the stock of currency in circulation grows at a lower rate than nominal GDP. In fact, figure 12 shows that when this effect is taken into account, currency in circulation remains essentially flat (hitting a minimum point in 2019) instead of increasing steadily as in figure 4. Since the path for the monetary base is still given by figure 10, from equation 9 we know that the lower path for currency in circulation must deliver a higher path for reserves. This is what figure 13 shows: The new path for reserves is higher than in figure 5. Given this path for reserves, SOMA interest expenses $i_t^R R_{t-1}$ are then obtained by multiplying them by the interest rates on reserves in figure 9 (which are assumed unchanged).
The resulting path, which is depicted in figure 14, is higher than in figure 6. However, the implications for remittances to the Treasury (constructed from equation 5) are small: Figure 15 shows that remittances are smaller than in figure 8, but the difference amounts to an accumulated total of only $23 billion between 2019 and 2022. Thus, redoing the calculations under a monetarist view does not raise any red flags in terms of remittances to the Treasury. However, the level of reserves by 2022 is about $360 billion in figure 13, compared with $100 billion in figure 5. Since the reserves threshold for resuming purchases of Treasury securities is $100 billion, the monetarist view indicates that the resumption point occurs later than in the benchmark calculations.

Something that is worth noting is the behavior of the money multiplier underlying these projections. Observe that the money multiplier is given by the ratio of money in circulation to the monetary base:

\[ \mu_t = \frac{M_t}{R_t + C_t} \]

Thus, in order to describe the behavior of the money multiplier during the projection period, I need to estimate the demand for money \( M_t \). An apparent difficulty in doing this is that the demand for money is usually considered to be highly unstable over time (for this reason, policy discussions are hardly ever conducted using the framework of a money demand function). In fact, figure A2 in the appendix, which
reproduces Lucas and Nicolini’s (2015) figure 2a, plots the inverse of the velocity of circulation of M1 versus the three-month Treasury bill at an annual frequency during 1915–2012 and shows that the demand for money was fairly stable during the 1915–1980 period, but that the relation broke down during 1981–2012. Since figure A1 shows that the demand for the currency component of M1 remained fairly stable, this means that the demand for demand deposits is what actually broke down. Figure A3 confirms this. Lucas and Nicolini (2015) argue that the reason for this breakdown is that since the appearance of money market funds in the early 1980s, some of the transactions that were previously done using checking accounts started being done with money market deposit accounts. In fact, appendix figure A4 shows that when money market deposit accounts are added to M1, the demand for money remains stable throughout the whole century.4

Motivated by this empirical evidence, I postulate a stable demand for money function of the following form:

\[ \frac{M_t}{PY_t} = A_m(i_t + \phi_m)^{\alpha_m}. \]

I then choose the parameters \( A_m, \phi_m, \) and \( \alpha_m \) to fit figure A4 in the appendix at the following three points:

<table>
<thead>
<tr>
<th>( i )</th>
<th>( \frac{M}{PY} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.039</td>
</tr>
<tr>
<td>0.05</td>
<td>0.023</td>
</tr>
<tr>
<td>0.135</td>
<td>0.14</td>
</tr>
</tbody>
</table>

The resulting values for \( A_m, \phi_m, \) and \( \alpha_m \) are 0.0341, 0.0592, and 0.8625, respectively. Figure 16 depicts the path for the money supply \( M \) implied by the demand for money function (equation 12). Dividing these numbers by the monetary base in figure 10 gives a path for the money multiplier \( \mu \) that is depicted in figure 17. This figure shows a striking result: After remaining fairly constant through 2019, the money multiplier is expected to almost double within a three-year period. Is it realistic to expect such a sharp increase over such a short period? The next section addresses this question.
On the money multiplier...

In the old days, short-term interest rates were positive but reserves did not earn interest. As a consequence, banks held just enough reserves to satisfy their reserve requirements. Defining \( A_t \) to be deposit accounts, the money multiplier was given by

\[
\mu_t = \frac{C_t + A_t}{C_t + R_t} = \frac{C_t + 1}{A_t + \rho_t},
\]

where

\[
\frac{C_t}{A_t} = \frac{C_t}{M_t - C_t} = \frac{1}{M_t - 1},
\]

and where \( \rho_t \) was the reserve requirement ratio. Using a reserve requirement ratio of 0.10 and the demand for money functions in equations 10 and 12 to evaluate the ratio \( M_t/C_t \), we get that the money multiplier \( \mu_t \) would have gone from 3.3 to 3.9 as the interest rate increased from 0 percent to 3 percent (the value by the end of our projection period). That is, in the old days, a doubling of the money multiplier under such an interest rate increase would have been far-fetched.

But we do not live in the old days, nor are we likely to go back to them. The current environment is characterized by abundant reserves and administered interest rates. In particular, the Fed puts a floor on safe short-term interest rates by paying interest on reserves and offering overnight reverse repos. In a context of abundant reserves, this would be enough to control short-term interest rates. But for the sake of argument, it may be useful to think that the Fed will also operate some type of lending facility that will put a ceiling on short-term interest rates. Moreover, for simplicity, assume that the floor and ceiling are the same, and therefore that the interest rate on reserves is equal to the interest rate on safe short-term assets. Observe that in this situation, banks should be completely indifferent between holding reserves and safe short-term assets.

In what follows, I argue that as a first approximation under such a policy regime, the size of the Fed’s balance sheet is completely irrelevant for equilibrium outcomes while the Fed gains full control of the money multiplier. To show this, let’s assume that the economy is in some equilibrium path for the price level \( P_t \), real GDP \( Y_t \), the interest rate \( i_t \), currency \( C_t \), deposits \( A_t \), and reserves \( R_t \). As a counterpart to their deposits \( A_t \), banks would generally hold safe short-term assets and reserves (among other assets). However, consider a first case in which banks hold no reserves at all (yes, I am abstracting from any reserve requirements) and only hold short-term Treasury securities in their portfolio of short-term assets. Observe that in this case, the banks receive interest payments from the U.S. Treasury and use these receipts to pay interest to their depositors (while the Fed stands idle on the side). Now consider an alternative scenario in which the Fed purchases all the holdings of short-term Treasury securities of the banks with reserves (increasing its balance sheet). In this scenario, the Fed is the one now receiving payments from the Treasury. The Fed uses these receipts to pay interest on reserves to the banks, which in turn use the receipts to pay interest to their depositors. But aside from having one more intermediary in the flow of payments ultimately going from the Treasury to depositors (and some differences on who is holding what), there is no substantial economic difference between the
two scenarios: Equilibrium outcomes should be exactly the same. Since the monetary base is higher in the second scenario but the quantity of money in circulation is exactly the same in both scenarios, the money multiplier is lower in the second scenario. However, this is completely irrelevant for equilibrium outcomes.

With respect to the question I asked at the end of the previous section, we can then say that a doubling of the money multiplier while the Fed shrinks the monetary base is a perfectly realistic outcome.

**Conclusion**

This article evaluated how sensitive projections of the net income of the Fed during the balance sheet normalization period could be to the incorporation of an explicit demand for money function. Previous projections assume that cash balances will grow at the same rate as nominal GDP. However, as interest rates increase, the demand for real cash balances will decrease, making cash grow at a lower rate than nominal GDP. As a consequence, reserves will decrease at a lower rate than expected and, given that the Fed pays interest on reserves, the net income of the Fed will be lower than expected. However, while all these effects are true, I find that the quantitative implications for the path of net income of the Fed are small.

A striking implication of the analysis is that the money multiplier is expected to double within a three-year period. However, I argue that this should pose no problem because under a system of administered interest rates that includes interest payments on reserves, banks are indifferent between holding reserves and safe short-term assets. Under such a monetary regime, the Fed directly controls the money multiplier by determining the size of its balance sheet.

A caveat of the analysis is that it was performed under a very specific path for the fed funds rate: the path obtained in the September 2016 Survey of Primary Dealers. It would be extremely interesting to perform the analysis under alternative scenarios. For instance, an interesting scenario would be one in which (maybe because of an inflation scare), interest rates rise quickly to much higher levels than in the benchmark case. In this scenario, the Fed would be paying much higher interest payments on reserves while reserves remain at higher levels because of a sharper drop in the demand for real cash balances, putting considerably more strain on the net income of the Fed than in the benchmark calculations. Another interesting scenario would consider alternative levels for the “normalized” size of the balance sheet, since this has not yet been decided by the FOMC (the $100 billion normalized reserves amount considered in this article was merely illustrative). However, analyzing such alternative scenarios would require taking a stance on what would be the associated paths for nominal GDP and the SOMA portfolio (which is highly sensitive to interest rates and the normalized size of the balance sheet). Determining such paths, while necessary for considering alternative scenarios, is outside the scope of this article.
APPENDIX: MONEY DEMAND EVIDENCE

This appendix reproduces the figures in Lucas and Nicolini (2015) that contain the money demand evidence used in this article.
1 A deferred asset would reflect the amount of future earnings that would be needed to be withheld to cover the Fed’s current operating losses. During the period of time that a deferred asset remains on its books, the Fed’s remittances to the Treasury would be equal to zero.

2 An exception is the New York Fed’s most recent report (Federal Reserve Bank of New York, 2018). Instead of assuming that currency will grow at the same rate as nominal GDP, this report’s benchmark scenario uses the median response to a question in the December 2017 Survey of Primary Dealers (SPD) about the expected level of currency in 2025. Currency is then assumed to grow at a constant growth rate during the projection period consistent with that 2025 level. The resulting annual growth rate of currency of 5 percent is actually larger than the 4 percent median annualized growth rate in nominal GDP between 2017 and 2025 implied by the December 2017 SPD expected paths for GDP and the PCE (personal consumption expenditures) deflator. That is, the most recent New York Fed report has moved even further away from the scenario explored in this article. For an argument supporting the view that currency may actually grow faster than nominal GDP due to foreign demand for U.S. currency, see Haasl, Paulson, and Schulhofer-Wohl (2018).

3 The money multiplier is regularly described in textbooks as the consequence of commercial banks being required to hold only a fraction of their deposits as reserves, lending the rest and thus creating additional money in circulation.

4 This monetary aggregate is known as money of zero maturity (MZM).

5 Under such a stable money multiplier, implementing the huge contraction in monetary base that the Fed is planning to do during the normalization period would have created extreme deflationary pressures.

REFERENCES


Marcelo Veracierto is a senior economist and research advisor in the Economic Research Department at the Federal Reserve Bank of Chicago.

© 2018 Federal Reserve Bank of Chicago

Economic Perspectives is published by the Economic Research Department of the Federal Reserve Bank of Chicago. The views expressed are the authors’ and do not necessarily reflect the views of the Federal Reserve Bank of Chicago or the Federal Reserve System.

Charles L. Evans, President; Daniel G. Sullivan, Executive Vice President and Director of Research; Anna L. Paulson, Senior Vice President and Associate Director of Research; Spencer Krane, Senior Vice President and Senior Research Advisor; Daniel Aaronson, Vice President, microeconomic policy research; Jonas D. M. Fisher, Vice President, macroeconomic policy research; Robert Cox, Vice President, markets team; Gene Amromin, Vice President, finance team; Leslie McGranahan, Vice President and Director, regional research; William A. Testa, Vice President, regional programs; Marcelo Veracierto, Senior Economist and Economics Editor; Helen Koshy and Han Y. Choi, Editors; Julia Baker, Production Editor; Sheila A. Mangler, Editorial Assistant.

Economic Perspectives articles may be reproduced in whole or in part, provided the articles are not reproduced or distributed for commercial gain and provided the source is appropriately credited. Prior written permission must be obtained for any other reproduction, distribution, republication, or creation of derivative works of Economic Perspectives articles. To request permission, please contact Helen Koshy, senior editor, at 312-322-5830 or email Helen.Koshy@chi.frb.org.

ISSN 0164-0682