Federal Reserve Bank of Cleveland Working Paper Series

Federal Reserve Balance-Sheet Policy in an Ample Reserves Framework: An Inventory Approach

Joseph G. Haubrich

Working Paper No. 23-25R

February 2026

Suggested citation: Haubrich, Joseph G. 2026. "Federal Reserve Balance-Sheet Policy in an Ample Reserves Framework: An Inventory Approach." Working Paper No. 23-25R. Federal Reserve Bank of Cleveland. https://doi.org/10.26509/frbc-wp-202325r.

Federal Reserve Bank of Cleveland Working Paper Series
ISSN: 2573-7953

Federal Reserve Balance-Sheet Policy in an Ample Reserves Framework: An Inventory Approach

Joseph G. Haubrich*

January 30, 2026

Abstract

Implementing monetary policy in an ample reserves regime means choosing a level of reserves that balances the cost of reserves against interest rate volatility. I use techniques from stochastic inventory theory to calibrate the size of the buffer needed to keep reserves above the ample level, and find the buffer size to be modest compared to the level of reserves needed to reach the ample level.

Keywords: Reserves, Monetary Policy
JEL codes: E58, D25

1 Introduction

In January 2019 the Federal Open Market Committee (FOMC) indicated it would implement monetary policy in an ample reserves regime "in which an ample supply of reserves ensures that control over the level of the federal funds rate and other short-term interest rates is exercised primarily through the setting of the Federal Reserve's administered rates, and in which active management of the supply of reserves is not required" (FOMC 2019). In May 2022, the FOMC further indicated that "the Committee intends to maintain securities holdings in amounts needed to implement monetary policy efficiently and effectively in its ample reserves regime" (FOMC 2022) even as it sought to reduce its securities holdings. Implementing this plan requires holding a buffer of securities so that shocks to demand and supply do not push reserves below the ample level. This paper calculates the buffer using techniques from stochastic inventory theory calibrated to the Federal Reserve's balance sheet and to the federal funds rate. Though different assumptions can at times lead to quite different buffers, in general my research suggests that the buffer can be relatively small compared to the amount needed for ample reserves, despite the FOMC taking

*Federal Reserve Bank of Cleveland, PO Box 6387, Cleveland, OH 44101-1387, 216 579 2802 jhaubrich@clev.frb.org. The views expressed here are solely those of the author and not necessarily those of the Federal Reserve Bank of Cleveland or the Board of Governors of the Federal Reserve System. The author has no other research support, financial relationships or potential conflicts of interest. In memory of Marvin Goodfriend, who both in person and through his work taught me about monetary regimes. Thanks to Matt Sobel for help on inventories, Julie Guess and Kim Zaikov for help with the data,Tom Phelan, Ed Knotek, Ricardo Lagos, Gaston Navarro, Micah DelVecchio, Jonathan Wright, Narayana Kocherlakota and participants at the WEAI Seattle, SNB Zurich and NBER Summer Institute Capital Markets and the Economy conferences for useful discussions, and Rachel Widra, Chris Healy and Ruchita Coomar for excellent research assistance.

actions consistent with a view that the costs of reserves falling below "ample" exceed the costs of having a large balance sheet.

Letting reseves get too low will move the fed funds rate off the floor, or even out of the target range, particularly if there are adverse supply and demand shocks, and while a larger balance sheet makes that less likely, it has costs of its own. The FOMC has explicitly stated the intention to keep the federal funds rate in the target range, as it made technical changes that were "intended to foster trading in the federal funds market at rates well within the FOMC's target range." It has acknowledged the costs to a large balance sheet both by reducing the balance sheet and by stating (FOMC 2014) that "the Committee intends that the Federal Reserve will, in the longer run, hold no more securities than necessary to implement monetary policy efficiently and effectively." The problem maps naturally into the newsvendor problem of stochastic inventory theory, which picks an inventory level that balances the costs of running out against higher storage costs. Here, it is the cost of an excessive fed funds rate against a large balance sheet. Individual banks face a similar problem with regard to their own reserves, and my results can be thought of as applying Poole's (1968) technique to the market as a whole. The solution from inventory theory chooses the optimal "critical fractile," the probability that the balance sheet gets too low (and the fed funds rate gets too high). Using a revealed preference approach, I note the fraction of time the FOMC has let the fed funds rate rise off the floor and then calibrate the buffer from the distribution of shocks to the Federal Reserve's balance sheet. I do this for both piecewise linear and non-linear demand for reserves.

My emphasis on the buffer needed to protect against demand and supply shocks stands in contrast to most recent work, which aims rather to quantify the level of reserves needed to reach ample. The Federal Reserve introduced the Senior Financial Officer Survey (SFOS) in 2018 to gauge banks' demand for reserves, and Fed staff analyses have used this to estimate when reserves would fall below ample and become scarce (Andros et al. 2019, Keating et al. 2019). More econometrically based work includes Langowski (2023), Lopez-Salido and Vissing-Jorgensen (2023), and Afonso et al., (AGLW, 2024). AGLW indeed begin their paper with "What level of central bank reserves satiates banks' demand for liquidity?"

Less work has considered the buffer needed to defend against shocks to demand and supply, and even that has often been qualitative. Gagnon and Sack (2019) argue that "the minimum level of reserves is conceptually murky, impossible to estimate, and likely to vary over time. The best approach is to steer well clear of it, especially since maintaining a higher level of reserves as a buffer has no meaningful cost." Lopez-Salido and Vissing-Jorgenson argue that "a buffer of several hundred billion dollars does not seem unreasonable given recent TGA volatility." Afonso et al. (2020) provide some qualitative results on the optimal buffer when the object is either to keep the fed funds rate in a given range or to minimize the probability of moving outside that range but they do not provide any estimates. Lagos and Navarro (2023) implicitly include a buffer in calculating their monetary policy confidence band, and a recently declassified FOMC memo by Schulhofer-Wohl and Zobel (2019) explicitly calculates a buffer. Neither takes the inventory approach of calculating the optimal stock-out probability.

This paper contributes to the recent but growing literature on implementing monetary policy with a large central bank balance sheet. Ihrig, Senyuz, and Weinbach (2020) provide a detailed description of the ample reserves approach, while Craig and Millington (2017) document changes in the federal funds market stemming from a large balance sheet. Afonso et al. (2020) provide a sophisticated theoretical and empirical justification for such a regime (that partly motivates this paper). Relative to them, I develop a simpler model of the reserves market but use techniques from inventory theory to quantify the optimal buffer more explicitly. Afonso et al. (2023) look at the optimal supply of central bank reserves under demand uncertainty. Relative to them I take a more inventory-theoretic approach, consider supply uncertainty and obtain some quantitative results. Early explanations of using administered rates as a tool of monetary policy include Goodfriend (2002) and Keister, Martin, and McAndrews (2008), who build on the early work on reserves markets of Poole (1968) and Frost (1971). Going back even further, Arrow (1958) remarks that Edgeworth (1888) initially developed stochastic inventory theory to study the Bank of England's balance sheet. AGLW (2024) estimate the demand for reserves and what constitutes an ample level, an issue also explored in Copeland, Duffie, and Yang (2025) and Afonso et al. (2022a). Reserve demand and interest rate control are discussed by Lopez-Salido and Vissing-Jorgensen (2023). Lagos and Navarro (2023) calibrate a dynamic equilbirum model of trading in the fed funds market and develop a monetary confidence band that lets the central bank meet the interest rate target with a specified degree of confidence. My aggregate approach is simpler and I think more transparent, and the method determines an optimal degree of confidence, which I then calibrate to FOMC behavior, rather than leaving it as an open choice.

The rest of the paper is structured as follows. Section 2 describes the basic model of the reserves market used in the rest of the paper. Section 3 applies stochastic inventory theory to the model of Section 2 and derives the optimal buffer stock of reserves. Section 4 calibrates the model to US data, namely, the federal funds rate and the Federal Reserve's balance sheet, first for the simple piecewise linear demand function, and then for non-linear demand with both demand and supply shocks. Section 5 concludes the paper.

2 A simple model

Understanding how demand and supply shocks interact with the size of the balance sheet to produce interest rate variability requires a more explicit model of the reserves market. This section develops a simple model of an ample reserves regime based on the important work of Afonso et al. (2020). Their approach incorporates the equilibrium of supply and demand by expressing the federal funds rate as a spread above a floor, here assumed to be interest on reserve balances (IORB), where the spread depends positively on the demand and negatively on the supply of reserves.[1]

$$rate = IORB + spread(reservesupply, reservedemand). \tag{1}$$

Equation (1) expresses the ability of the central bank to move the target rate by changing the administered rate $IORB$, and how shocks to demand and supply can cause fluctuations around that rate.

In an ample, or floor, regime, the central bank wishes to minimize the variation in the spread subject to balance-sheet costs, choosing a balance-sheet target T to minimize the expected loss

$$\min_T \mathrm{E}_{s,\delta}\{L[spread(T; s, \delta), T + s]\}. \tag{2}$$

where L is the loss function, E denotes the expected value, and s and δ are the supply and demand shocks.

[1] Richer equilibrium approaches to the federal funds market can be found in Hamilton (2020), Afonso, Armenter, and Lester (2019), Bianchi and Bigio (2022), and Lagos and Navarro (2023).

I initially specialize the form of the spread to the following inverse demand function

$$spread = Max[D - a(T + s - \frac{\delta}{a}), 0]. \tag{3}$$

where D is the y-intercept of the inverse demand curve; T is the (target) level of reserves; δ is the demand shock, modeled as a parallel shift in the demand curve shifting the intercept by the amount δ; s is the supply shock; and a is the slope of the inverse demand curve. This captures, in a simple way, the main features of the current reserves market: a demand for reserves that slopes downward until it reaches a floor at the rate of interest on reserves, at which point any amount supplied will be willingly held. It puts the demand and supply shocks on the same footing.

This formulation, following Goodfriend (2002), puts a kink point in the demand function at what we will also call the "saturation point" of reserves: adding more reserves beyond this point does not lower the spread: demand is saturated. An ample reserves framework puts reserves above this saturation point, and policy moves the fed funds rate by changing the IORB.

This formulation makes several modeling choices, which have several pros and cons. As in Afonso et al. (2020) and Lopez-Salido and Vissing-Jorgensen (2023), the relevant price is not the federal funds rate *per se* but the spread, so that an increase in the IORB shifts up the demand curve. This captures an essential element of how such floor systems are supposed to work. It also assumes that the IORB, the interest on reserve balances, functions as the floor, which ignores the often complicated relationships between the IORB and other administered rates such as the overnight reverse repo (ONRRP) rate. (Lopez-Salido and Vissing-Jorgensen explore this issue in detail.) Recently, the ONRRP rate and the IORB have been 5 and 15 basis points above the bottom of the target fed funds rate range established by the FOMC, though the ONRRP spread was reduced to zero in December 2024. Likewise, this formulation treats the other portions of the Fed's balance sheet as exogenous factors: as of October 22, 2025, while the System held \$2.9 trillion of reserves, currency in circulation was \$2.4 trillion and reverse repurchase agreements were \$5 billion. In monetary policy discussions it is tempting to equate the Fed's balance sheet with the level of reserves, but they are by no means identical. The demand curve is piecewise linear, a simplification relative to Afonso et al. (2020) (but similar to Afonso et al. 2023), and although less realistic, allows for a cleaner calibration in the inventory context. The formulation also assumes that the aggregate level of reserves matters for the fed funds rate, whereas evidence suggests that the distribution of reserves among banks matters as well (Copeland, Duffie, and Yang, 2025). See Ihrig, Senyuz, and Weinbach (2020) for a more detailed description of the ample reserves regime, and Mester (2024) for a nice description of subtle aspects of the operating framework.

The question of rate control is then about how to respond to the shocks that move the spread, and how much tolerance the Committee has for deviations from the target. Again, this takes as given a specific target for the federal funds rate. The monetary policy question of the appropriate target fed funds rate is separate from the level of reserves needed for efficient and effective implementation of that policy. Likewise I take a floor system as given and abstract from questions about floor versus corridor systems. See Arce et al. (2020) for a representative contribution.

3 Inventory

The above model of the reserve market illustrates the interaction between reserve demand, reserve supply via the Federal Reserve's balance sheet, and the level of interest rates. Shocks to demand and supply create interest rate variability when reserves (and the balance sheet) are low. The optimal balance sheet trades off the costs of interest rate variability against the costs of a larger balance sheet, and this maps quite naturally into a stochastic inventory problem, where a vendor balances the costs of excessive inventory against the costs of running out.

Poole (1968) considered an individual bank's demand for reserves subject to shocks, and his analysis can be applied to the central bank. The central bank aims to have a balance sheet no larger than necessary to control the fed funds rate. I interpret this as, without shocks, the balance sheet should be at the smallest level that puts the FFR at the IORB floor (in the Afonso et al. (2020) model, that would be a spread of zero). Without loss of generality, I label the combined demand and supply shocks $(s - \frac{\delta}{a})$ as a supply shock that adds s to the balance sheet, with a cumulative distribution $F(s)$ and density $f(s)$ with mean μ. If the balance sheet drops below the minimum level and the spread rises above zero, there is a per dollar penalty cost c_p of the deficiency (which might be reputational, of course). With a linear demand for reserves, as in Section 2, this is equivalent to assigning a penalty to spreads above zero but is also compatible with a non-linear demand for reserves, as long as the cost of reserve shortfalls is linear in quantities. Holding a large balance sheet has a per dollar cost c_B for assets on the balance sheet. An alternative model might follow the inventory approach more slavishly and only assign costs to balance sheets above the minimum level of ample reserves, but the results are quite similar.[2] The piecewise linear cost structure follows the inventory literature, but Benartzi and Thaler (1995), in a loss aversion context, find that the specific functional form is not critical for practical results. Afonso et al.(2020) also assume linear costs in their theoretical model, while Afonso et al. (2023) consider non-linear costs in a theoretical context.

3.1 One-period case

Let T (for target) denote the amount of reserves after the open market operation. Let A denote the minimum level of the reserves needed for an ample balance sheet, the minimum level where the fed funds rate equals the IORB floor and the spread is zero. (In the model of Section 2, $A = \frac{D}{a}$.) Then, following Poole (1968), the (one-period) holding and shortage cost is

$$L(T) = \int_{-\infty}^{\infty} c_B(T+s)f(s)ds + \int_{-\infty}^{A-T} c_p(A-T-s)f(s)ds. \tag{4}$$

Given reserves of size T, the rate rises above the floor if the shock is negative enough to drive reserves below A, that is, if $T + s < A$ or equivalently $s < A - T$, making the penalty cost $c_p(A - T - s)$, so that the loss increases as the fed funds rate gets further above the floor. Given a shock s, the total balance sheet is $T + s$, resulting in a cost of $c_B(T + s)$.[3] The objective is to minimize the expected cost $L(T)$.[4]

[2] For a discussion of the costs of having a large balance sheet related to political economy and credit allocation, see Copeland, Duffie, and Yang (2025).

[3] This embeds the not entirely satisfactory assumption that for extreme negative values of s, the balance sheet is negative and the balance-sheet holding cost becomes negative. This has an aspect of realism, as under a corridor system there is a borrowing as well as a lending rate. But in any case, I judge the probability to be small and empirically irrelevant. Section 6.4 in the Appendix works out the case of a lower limit, say, for a distribution with finite support, such as a beta distribution.

[4] An alternative class of loss functions, discussed by Afonso et al. (2020), has the social planner minimizing the probability of intervening in the market (preventing the spread from exceeding some tolerance) plus some cost of having a large balance sheet. Section 6.1 in the Appendix considers that approach.

Using Leibniz's rule (Boas, 1966) the first-order condition becomes

$$\frac{dL}{dT} = c_B - c_p \int_{-\infty}^{A-T} f(s)ds = 0. \tag{5}$$

Noting that $\int_{-\infty}^{A-T} f(s)ds = F(A-T)$ the optimal balance-sheet level T^* is chosen so that

$$F(A - T^*) = \frac{c_B}{c_p} \tag{6}$$

where $\frac{c_B}{c_p}$ is the *critical fractile* and gives the optimal probability of letting the fed funds rate rise above the floor. Conversely, the optimal level is

$$A - T^* = F^{-1}(\frac{c_B}{c_p}) \tag{7}$$

The reverse of this, $B = T^* - A = -F^{-1}(\frac{c_B}{c_p})$, gives the optimal "buffer" level of reserves above the point of transition to reserve scarcity.

Intuitively, $F(A-T^*)$ has the properties we expect. As c_B increases, the cost of holding a larger balance sheet increases, $F(A-T^*)$ increases, and T^* falls. Provided $c_B \geq 0$, an increase in the penalty cost c_p decreases $F(A-T^*)$ and so increases T^*; as the cost of letting the fed funds rate rise above its target increases, the balance sheet increases to reduce that probability.

The critical fractile is the key component of the analysis, and can be generalized beyond the one-period model used here. Appendix 6.3 derives the fractile for a discrete-time dynamic model with a cost of acquiring reserves and set-up costs.

4 Application: Calibrating relative costs and optimal buffer

The theory in Section 3 interprets the point at which reserves become scarce and the effective fed funds rate rises above the floor as similar to the inventory concept of stock-out, where inventory hits zero. The optimal level of inventory, determined by the critical fractile, is then the optimal buffer stock for keeping reserves in the ample regime. Taking a revealed preference approach, this section calculates the optimal buffer using data on bank reserves, the Fed's balance sheet, and the fed funds market. The critical fractile, equation (6), is calibrated as the fraction of time that the fed funds rate rises above the floor. Given the critical fractile, the next step is to estimate the distribution of supply shocks. Inverting that distribution at the critical fractile implies a value for the optimal buffer, $F^{-1}(F(B)) = B$. As a robustness check, I use several definitions of what it means for the fed funds rate to be off the floor, and several estimates of the distribution of shocks, leading to a range of buffer values.

4.1 Data

There are several ways to determine when the fed funds rate is off the floor and reserves are no longer ample. The one that best matches the model of Section 2 compares the effective fed funds rate with the interest on reserves. Using this definiton, the floor is the interest rate on excess reserves, IOER, from October 15, 2008, until it is replaced by the interest on reserve balances on July 29, 2021, which I use until July 5, 2023. Both series are taken from the Federal Reserve's

Data Download site.[5] The effective fed funds rate is calculated by the Federal Reserve Bank of New York "as a volume-weighted median of overnight federal funds transactions reported in the FR 2420 Report of Selected Money Market Rates" (FRB NY, n.d.). A more conservative approach (in the sense of less time above the floor) would consider when the EFFR rises above the target range established by the FOMC. The target range for the fed funds rate starts December 16, 2008, and so starts later than the payment of interest on reserves. This is taken from the Federal Reserve via FRED: DFF Federal Funds Effective Rate, Percent, Daily, Not Seasonally Adjusted, DFEDTARU Federal Funds Target Range - Upper Limit, Percent, Daily, Not Seasonally Adjusted.

Over most of the interest on reserves period, however, macroeconomic concerns (financial crises, COVID) dominated decisions about the balance sheet. The period where the FOMC appeared to be concerned about keeping the fed funds in range was much shorter. What the FOMC termed balance-sheet normalization commenced in October 2017 and ended in August 2019, two months ahead of the originally scheduled stop. Over that time the spread between the interest rate on reserves and the target fed funds range was adjusted four times. The Implementation Note in each case explained that the change was "intended to foster trading in the federal funds market at rates well within the FOMC's target range." The Federal Reserve Bank of New York's *Liberty Street Economics Blog* (Afonso et al. 2022b) notes that these technical adjustments were "a tool the Fed can deploy to keep the FOMC's policy rate well within the target range." Thus, for purposes of computing the critical fractile, I use the period from January 2018 to March 2020.

Lopez-Salido and Vissing-Jorgensen (2023) likewise consider the 2018-2019 period as particularly "instructive" (p.27) for assessing their model of interest rate control. Lagos and Navarro (2023) calibrate their model to fed funds trading in 2019. Clearly there is a degree of arbitrariness in choosing these dates; if we consider the most recent episode of shrinking the balance sheet beginnng in June 2023, there would be an additional 1100 days where the EFFR did not rise above the IORB, though again, this was starting from a point where the FOMC was acknowledging that the balance sheet and reserves were higher than desired.

As explained above, the critical fractile is the fraction of time that the federal funds rate rises above the floor. The EFFR is usually below the IORB, indicating ample reserves. However, over the January 2018 to March 2020 period, when rate control issues were more salient, the EFFR exceeded the IORB 29.9 percent of the time. This pins down the critical fractile of stocking out $F(A - T^*) = \frac{c_B}{c_p}$ from equation (6) to 29.9 percent.

The second step in calibrating the buffer involves estimating $F(s)$, the distribution of supply shocks to the Fed's balance sheet. I first follow Lagos and Navarro (2023) and look directly at daily changes in reserves during the ample reserves regime, and then daily changes in an important autonomous factor, the Treasury General Account (TGA) ($0.4 trillion), whose level is controlled by the Department of the Treasury. If the Treasury spends money, drawing down the TGA, reserves increase. The publicly available balance-sheet data are weekly, but I use confidential daily data from the Federal Reserve FR-34 report, from December 1, 2015 to July 10, 2024. To preserve confidentiality I do not report the summary statistics for the daily data. For weekly data, the

[5] `https://www.federalreserve.gov/datadownload/Choose.aspx?rel=PRates` The Federal Reserve has paid interest on reserve balances since October 6, 2008. Initially there were two rates: one for required reserves (interest on required reserves, IORR), and one for excess reserves (interest on excess reserves, IOER). Reserve requirements became zero effective March 24, 2020. Not needing to distinguish the rates, the Board of Governors established the interest rate on reserve balances effective July 29, 2021. Since its inception, the IORB has been set at 15 basis points above the lower limit of the federal funds target range. The IORR stood at 25 bp above the bottom of the range from the end of 2008 to June of 2018, when it decreased to 20 bp, after which it fluctuated slightly, reaching a low of 5 bp in 2019 before rising to 15 again in the summer of 2021, just before it was replaced by the IORB.

mean change in reserves is $1.8 billion with a standard deviation of $82 billion. For changes in the Treasury General Account the mean is $1.2 billion with a standard deviation of $53 billion.

4.2 Buffer size

Given an observed critical fractile, equation (6), the next step in calculating the optimal buffer requires estimating the distribution of balance-sheet shocks, $F(s)$. A normal distribution is the usual place to start (Porteus, 2002, p. 12) but it may not be the best choice. A Shapiro-Wilks test and a QQ plot (not shown) indicate substantial non-normality and fat tails. I estimate the density using a Gaussian kernel (implemented in R).[6] Using different kernels or bandwidth had minimal impact on the estimation. Since kernel estimates can lead to estimates above the range of the data (Härdle 1991 section 3.0), I also compute quantiles on the actual empirical distributions for comparison.[7] Using the kernel estimate, the critical fractile of 29.9 percent implies an optimal buffer of $16.1 billion, while the empirical distribution implies a smaller buffer of $15.4 billion. Recall that these are numbers for the buffer, and total reserves should be grossed up by the reserves needed to attain the ample level.

The above procedure assumes that penalty costs start when the effective fed funds rate rises above the IORB rate. A more conservative approach might judge the fed funds rate as excessive when the effective fed funds rate rises above the upper limit of the target range. This happened only once in the January 2018 to March 2020 period, implying a critical fractile of 0.18 percent, in turn implying a ratio of penalty costs to balance-sheet costs of 555, and an optimal buffer size of $108 billion under the estimated distribution.

Table 1 collects the calibrated buffers according to the various criteria for the January 2018 to March 2020 time period. The first panel uses the changes in reserves as the supply shock, estimated via a kernel and the empirical distribution. The second panel uses the daily changes in the TGA as the supply shocks.

Table 1: Optimal Buffer Levels Under Different Assumptions

Criterion	Distribution	Sample	Critical Fractile	Buffer($B)
Changes in reserves				
EFFR >IORB	Kernel	2018:01-2020:03	29.9%	16.1
EFFR> Upper Target	Kernel		0.18%	108.3
EFFR >IORB	Empirical		29.9%	15.4
EFFR> Upper Target	Empirical		0.18%	250.0
TGA changes				
EFFR >IORB	Kernel	2018:01-2020:03	29.9%	8.2
EFFR> Upper Target	Kernel		0.18%	88.9
EFFR >IORB	Empirical		29.9%	8.1
EFFR> Upper Target	Empirical		0.18%	140.7

Buffers in billions of dollars. Source: Author's calculations

[6] Using the R command density() with the default kernel (Gaussian) and bandwidth (nrd0).

[7] In their theoretical example, Afonso et al. (2020) use a uniform distribution.

The buffers differ depending on data and estimation technique, but they show a similar pattern. Total reserves are subject to more shocks than just changes in the TGA; so offsetting them requires a larger buffer for the balance sheet. Avoiding a one-in-500 downside shock requires a larger buffer than a one-in-three shock; so protecting against a breach of the fed funds target range requires a larger buffer than keeping rates below the interest on reserve balances. The kernel density apparently misses some of the tail weight seen in the empirical distribution and so implies a smaller buffer to protect against the rare shock that pushes rates outside of the FOMC's target range.

How does the estimated buffer size compare with the amount of reserves needed to be ample (an estimate of A above)? Waller (2022) noted that"financial markets worked well" when bank reserves were about 8 percent of GDP, which, as of 2023 Q2, equals about \$2.09 trillion, equivalent to a SOMA size of about \$5.14 trillion at the July 2023 SOMA to reserves ratio (July 6, 2023 H.4.1). This makes the optimal buffer about 6 percent of reserves or 2.5 percent of the SOMA portfolio. AGLW (2024) estimate that the transition between scarce and ample reserves happens at about 8 percent of bank assets, which, as of February 2023, stood at \$22,895 billion (Federal Reserve H.8), suggesting reserves of \$1,832 billion or a SOMA level of \$4,487 billion. Either number suggests that the buffer size is relatively small compared to the minimal level required for ample reserves.

As previously mentioned, how frequently the fed funds rate exceeds the interest on reserves depends on the sample, and this adds an arbitrary element to the buffer calculation. Figure 1 lets the reader judge the effect of different sample sizes by plotting the buffer size (in billions of dollars) for each critical fractile. Note that the fractile must be considerably reduced before the buffer even approaches \$100 billion, which is still a small fraction of the reserves needed to stay ample.

In contrast to the estimates for ample or abundant reserves, the previous literature has had little direct discussion of buffer size. An earlier version of Lopez-Salido and Vissing-Jorgensen (2023, p.25) remarks that "A buffer of several hundred billion dollars does not seem unreasonable given recent TGA volatility." A recently declassified FOMC memo by Schulhofer-Wohl and Zobel (2019) calculates a buffer of \$190 billion on estimated reserve demand of \$860 billion, with the buffer based on a 95 percent confidence interval about estimated reserve demand, whereas Table 1 is based only on supply shocks. Lagos and Navarro (2023) estimate a 95 percent confidence interval for supply shocks of [-\$115 billion,\$99 billion]; my estimated confidential interval is [-\$66 billion,\$70 billion], somewhat smaller. Table 1 does not include demand uncertainty, suggesting that the estimates may be on the small side, but have the advantage of calculating a buffer without needing to estimate the level of ample reserves. A back-of-the-envelope approach would be to add the buffers of Table 1 to any estimated reserve demand.

4.3 The critical fractile and relative costs

Since the critical fractile is determined by a ratio of costs, a given critical fractile implies a cost ratio. Then $\frac{c_B}{c_p} = 0.299$ or $c_p = 3.3c_B$. This is consistent with the FOMC acting as if a dollar of reserve deficiency is a bit more than 3 times as costly as maintaining an extra dollar on the balance sheet. FOMC statements might suggest at least an ordinal ranking, as they first discuss the fed funds rate and then the balance sheet.

We can take this one step further. There are few, if any, estimates of the direct costs of a large balance sheet, which include banks holding higher capital against reserves, reduced bank participation in capital markets and a soaking up of scarce safe collateral (Schnabel 2024), all of which are hard to estimate. There is also a direct fiscal cost. Lucas (2022) notes that between

2008 and 2019, the IORB exceeded the rate on 3-month Treasury bills by 15 basis points. (As of October 8, 2025, the IORB rate was 4.15 percent and the Treasury Constant Maturity 3-month T-bill rate was 4.01 percent, per the H.15 report.) That represents a cost of \$1.5 million per \$1 billion of excess balance sheet, which, using the 3.3 ratio calculated above, implies a fiscal cost of reserves dropping below the ample level by \$1 billion to be \$5.0 million.

4.4 Non-linear demand

As mentioned above, the piecewise linear demand function assumed here is a simplification. The revealed preference calibration approach, however, works with a non-linear demand curve. Given a demand curve and quantity of reserves, shocks to supply and demand will create a distribution of interest rates, and assuming optimization by the monetary authority, this will imply a critical fractile and inverting the distribution will provide the optimal buffer. The general non-linear theory will still imply a fractile that can be used to calibrate the model.

In this paper I do not estimate the demand for reserves and instead take estimates from AGLW (2024). To account for parameter uncertainty, I calculate buffers using two sets of parameters from AGLW. The first uses their initial parameter choices, based on historical spread and reserves data. The second set uses their estimated values, the result of a non-linear contrained optimization. The form of their demand function is

$$spread = p^* + (\arctan(\frac{\theta_1 - q + q*}{\theta_2}) + \frac{\pi}{2})\theta_3. \tag{8}$$

spread and p^* are in basis points, and q and $q*$ are in percent of aggregate bank assets, which in 2018-2020 was about \$18 trillion. Table 2 lists the parameters taken from AGLW used in the calibrations. It also notes that the threshold between scarce and ample reserves that AGLW use is $\theta_1 + \theta_2/\sqrt{3}$, the point where the slope of the demand curve changes the fastest (a second derivative condition). However, the second derivative criterion used by AGLW results in a spread below zero at the ample level. This would often result in a negative buffer, since to match the critical fractile requires reserves below what AGLW define as ample. It seems more appropriate to define ample as the point where the spread is zero, being more in line with the FOMC's May 2022 press release on Plans for Reducing the Size of the Federal Reserve's Balance Sheet, which stated that it planned to "stop the decline in the size of the balance sheet when reserve balances are somewhat above the level it judges to be consistent with ample reserves." The buffer calculated by the calibration will be relative to this point. The table also lists the AGLW ample level, for comparison.

As before, the buffer is calibrated by matching the critical fractile, a revealed preference that assumes the FOMC is optimizing, even if the costs are not observable to the econometrician. Given a target level of reserves, the distribution of supply shocks implies a distribution of the spread (fed funds rate less IORB). The calibration chooses the target level of reserves so that the amount of time the fed funds rate exceeds the IORB matches the data: that is, matches the observed critical fractile. The buffer is then the difference between the target level of reserves and the ample level of reserves defined as having a zero spread. This is done using data on reserve shocks and TGA shocks, for both the initial parameters in the AGLW paper and their final estimated parameters. THat is, the buffer is the difference between the target level and level considered ample.

A non-linear demand curve requires taking demand shocks more seriously. To account for demand shocks, I allow shocks to the p^* term in equation 8. This adds a vertical shift to the demand curve. I produce the shocks by taking 2000 draws from a normal distribution with mean -18.75 and

Table 2: Parameters for the AGLW Demand Function

Parameters	Initial	Estimated
θ_1	7.3	4.82
θ_2	2.89	5.55
θ_3	15.3	18.22
$q*$	0.0	0
$p*$	-20	-20
Implied level of Ample	8.97	8.02
Ample level, $ billions	$1614.3	$1444.4
Zero-point Ample	8.97	8.02
Zero-point Ample level, $ billions	$1614.3	$1444.4

Units: percentage of total bank assets, $p*$ in basis points.
Source: Author's calculations

standard deviation 2, meant to approximate the distribution of p^* in Table 6 of AGLW. Combined with supply shocks, this produces 4 million equilibrium spreads. As before, the target level of reserves is chosen to match the fractile of time that the spread rises above zero (alternatively, when the EFFR rises above the top of the FOMC's range) and the buffer is calculated as the difference between the target level and the ample level. Table 3 reports the results. The demand curve is estimated with uncertainty, and as a simple way to account for parameter uncertainty, Table 3 shows results for both the initial and estimated parameters. This is not meant to, and does not fully replicate the statistics of the fed funds market, but rather indicates how demand and supply shocks interact in the setting of the buffer.

The values in Table 3 are more varied than the values for the piecewise linear demand in Table 1. Partly this is because using estimated parameters leads to buffers about twice as high as using the initial parameters. The estimated values are to be preferred, but the differences emphasize the uncertainty around the buffers. The extreme non-linearity of the demand function also creates some differences. One result is that the optimal target level is often below the level AGLW define as ample, because in the AGLW demand specification the FFR-IORB spread is negative at the level of ample reserves. Given the observed shocks, sometimes matching the critical fractile implies that the target level should be *below* the ample level. Though perhaps counter-intuitive, this matches the theoretical results in Afonso et al. (2020). Excluding demand shocks, the buffers would be predominately negative or near zero, as reported in Table 4, which shows the target level of reserves and required buffer if only estimated supply shocks exist.

A comparison of the numbers in Tables 1, 3, and 4 can help parse out the important differences in the buffer calculations, and points out that reasonable changes in assumptions can lead to large changes in optimal reserves. All three use the same distribution of supply shocks. Consider the buffer needed to keep the fed funds rate below the IORB. For the calibrations without demand shocks the buffers in Table 4 are much smaller than those in Table 1, indicating the difference that using non-linear demand makes. For the two calibrations with non-linear demand, Table 4 has smaller buffer values than Table 3, indicating that adding demand shocks adds uncertainty and requires a bigger buffer. Despite this, I would argue that the best estimate of buffer values–using the estimated parameters from AGLW with demand shocks–are in the $85-170 billion range.

Table 3: Optimal Buffer Levels with AGLW demand, With DEMAND shocks, Kernel estimates

Criterion	Target	$Target	Buffer	$Buffer
Changes in Reserves				
Initial Values				
AGLW Ample	8.968	1614.3		
Zero Ample	8.08	1454.4		
EFFR >IORB	8.57801	1538.6	0.498	84.2
EFFR> Upper Target	7.50	1350	-0.58	-104.4
99% VaR EFFR >IORB	9.4652	1702	1.385	246.6
99% VaR EFFR> Upper Target	7.3	1314	0.78	-140.4
Estimated Demand Values				
AGLW Ample	8.024294	1444.37		
Zero Ample	7.66	1378.9		
EFFR >IORB	8.59	1547.1	0.93	168.2
EFFR> Upper Target	6.45	1161	-1.21	-217.9
99% VaR EFFR >IORB	10.39	1870.2	2.73	491.3
99% VaR EFFR> Upper Target	6.2	1116	-1.46	-262.9
TGA changes				
Initial Values				
AGLW Ample	8.968	1614.3		
Zero Ample	8.08	1454.4		
EFFR >IORB	8.577	1543.8	0.497	89.4
EFFR> Upper Target	7.5	1350	-0.58	-104.4
99% VaR EFFR >IORB	9.5	1710	1.42	255.6
99% VaR EFFR> Upper Target	7.243	1303.74	-0.837	-150.66
Estimated Demand Values				
AGLW Ample	8.024294	1444.37		
Zero Ample	7.66	1378.9		
EFFR >IORB	8.6	1548	0.94	169.1
EFFR> Upper Target	6.4	1152	-1.26	-226.9
99% VaR EFFR >IORB	10.54	1897.2	2.88	518.3
99% VaR EFFR> Upper Target	6.2	1116	-1.46	-262.9

Target and Buffer expressed as percent of aggregate bank assets. $Target and $Buffer in billion dollars. Buffers relative to the level needed to set the spread to zero (termed Zero Ample). Source: Author's calculations

Table 4: Optimal Buffer Levels with AGLW demand, No demand shocks, Kernel estimates

Criterion	Target	$Target	Buffer	$Buffer
Changes in Reserves				
Initial Values				
AGLW Ample	8.968	1614.3		
Zero Ample	8.08	1454.4		
EFFR >IORB	8.08087	1454.56	0.001	0.16
EFFR> Upper Target	6.125	1102.5	-1.955	-351.9
99% VaR EFFR >IORB	8.087	1455.7	0.007	1.3
99% VaR EFFR> Upper Target	6.1179	1101.2	-1.96	-353.2
Estimated Demand Values				
AGLW Ample	8.024294	1444.37		
Zero Ample	7.66	1378.9		
EFFR >IORB	7.6617	1379.1	0.002	0.2
EFFR> Upper Target	4.41	793.8	-3.25	-585.1
99% VaR EFFR >IORB	7.667	1380.1	0.007	1.2
99% VaR EFFR> Upper Target	4.405	792.9	-3.255	-586
TGA changes				
Initial Values				
AGLW Ample	8.968	1614.3		
Zero Ample	8.08	1454.4		
EFFR >IORB	8.08044	1454.5	0.0004	0.1
EFFR> Upper Target	6.119	1101.42	-1.961	-352.98
99% VaR EFFR >IORB	8.086	1455.48	0.006	1.08
99% VaR EFFR> Upper Target	6.1179	1101.22	-1.9621	-353.18
Estimated Demand Values				
AGLW Ample	8.024294	1444.37		
Zero Ample	7.66	1378.9		
EFFR >IORB	7.66137	1379.05	0.0014	0.15
EFFR> Upper Target	4.4065	793.17	-3.2535	-585.73
99% VaR EFFR >IORB	7.67	1380.6	0.01	1.7
99% VaR EFFR> Upper Target	4.404	792.72	-3.256	-586.18

Target and Buffer expressed as percent of aggregate bank assets. $Target and $Buffer in billions of dollars. Buffers relative to the level needed to set the spread to zero (termed Zero Ample). Source: Author's calculations

The more problematic numbers are the negative buffers to keep the fed funds rate in the FOMC's target range. This is an interaction between the tail of the distribution and the non-linearity of the demand curve. It does suggest, however, that the more conservative buffer is the one for the IORB, and perhaps the infrequent breaching of the top of the target suggests that the occurrence was a mistake and not the best revelation of Committee preferences.

Comparing Table 3's estimates of ample with other estimates in the literature provides an additional check on the calibration. Ample reserves in Table 3 range from 7.7 to 8.9 percent of bank assets, or between $1380 and $1614 billion, a bit below the $1830 derived from Governor Waller's speech. Langowski (2023) finds the ample level at adjusted reserves of 7-11 percent of GDP, or between $1957 and $3070 billion (based on 2023 Q2 GDP). The 2019 Schulhofer-Wohl and Zobel memo uses $860 billion as the estimated ample level, and Andros et al. (2019) calculate a lowest comfortable level of reserves at between $900 billion and $1.5 trillion. The different definitions and parameter values in Table 3 also highlight the uncertainty surrounding estimates of ample, and a cautious central banker might want to consider the $236 billion between the high and low values as justification for additional buffers.

The most direct comparison with the "Target" levels of Table 3, which include both ample reserves and a buffer, is the monetary confidence band of Lagos and Navarro (2023), which calculates that $670 billion of reserves would keep the fed funds rate in within 25 basis points of the IOR rate with 95 percent confidence, and $850 billion would provide 99 percent confidence. Table 3 also lists the 99 percent value-at-risk levels for the spread, that is, the target level of reserves needed to make sure the spread stays above zero 99 percent of the time, or stays below the top of the fed funds target range 99 percent of the time. The 99 percent value is a bit higher than that found by Lagos and Navarro, coming in at between $1700 and $1190 billion, a function of the higher level of ample reserves coming from the AGLW model.

5 Conclusion

The problem of keeping reserves at a level no larger than needed for effective and efficient interest rate control maps naturally into a stochastic inventory problem, trading off balance-sheet costs against interest rate variability. The inventory buffer based on revealed preference has several advantages over previous work. It allows the buffer to be calibrated separately from the calculation of the ample level of reserves. By revealed preference, it endogenizes the critical fractile, accounting for how the FOMC has balanced the costs without assuming a 99 or 95 percent confidence interval. Assuming a piecewise linear demand for reserves, the optimal buffer is modest relative to the balance sheet required to maintain an ample regime. A perhaps more realistic non-linear demand for reserves leads to a higher though still modest optimal buffer. This has implications for the optimal size of the Fed's balance sheet and therefore the allowable level of quantitative tightening (QT).

References

Afonso, Gara, Roc Armenter, and Benjamin Lester. 2019. "A Model of the Federal Funds Market: Yesterday, Today, and Tomorrow." *Review of Economic Dynamics, Fragmented Financial Markets*, 33 (July): 177–204. https://doi.org/10.1016/j.red.2019.04.004.

Afonso, Gara, Darrell Duffie, Lorenzo Rigon, and Hyun Song Shin. 2022a. "How Abundant Are Reserves? Evidence from the Wholesale Payment System." Staff Report Number 1040. Federal Reserve Bank of New York. https://www.newyorkfed.org/research/staff_reports/sr1040.html.

Afonso, Gara, Gabriele La Spada,Thomas M. Mertens and John C. Williams. 2023. "The Optimal Supply of Central Bank Reserves under Uncertainty."FRB New York Staff Reports no. 1077, November 2023, revised December 2023. https://doi.org/10.59576/sr.1077

Afonso, Gara, Domenico Giannone, Gabriele La Spada, and John C. Williams. (AGLW) 2024. "Scarce, Abundant, or Ample? A Time-Varying Model of the Reserve Demand Curve." Staff Report Number 1019. Federal Reserve Bank of New York. https://www.newyorkfed.org/research/staff_reports/sr1019.html.

Afonso, Gara, Kyungmin Kim, Antoine Martin, Ed Nosal, Simon Potter, and Sam Schulhofer-Wohl. 2020. "Monetary Policy Implementation With an Ample Supply of Reserves." Working Paper 2020–02. Federal Reserve Bank of Chicago. https://www.chicagofed.org/publications/working-papers/2020/2020-02. https:/doi.org/10.21033/wp-2020-02

Afonso ,Gara, Lorie Logan, Antoine Martin, William Riordan, and Patricia Zobel. 2022b. "How the Fed Adjusts the Fed Funds Rate within Its Target Range," Federal Reserve Bank of New York Liberty Street Economics, January 12, 2022, https://libertystreeteconomics.newyorkfed.org/2022/01/how-the-fed-adjusts-the-fed-funds-rate-within-its-target-range/.

Andros, Joseph, Michael Beall, Francis Martinez, Tony Rodrigues, Mary-Frances Styczynski, and Alex Thorp. 2019. "Approaches to Estimating Aggregate Demand for Reserve Balances." FEDS Notes. Board of Governors of the Federal Reserve System. https://doi.org/10.17016/2380-7172.2459.

Arce, Óscar, Galo Nuño, Dominik Thaler, Carlos Thomas, 2020 "A large central bank balance sheet? Floor vs corridor systems in a New Keynesian environment." *Journal of Monetary Economics*, 114, 2020, Pages 350-367, ISSN 0304-3932, https://doi.org/10.1016/j.jmoneco.2019.05.001.

Arrow, Kenneth J. 1958. "Historical Background." In *Studies in the Mathematical Theory of Inventory and Production*, edited by Samuel Karlin, Herbert E. Scarf, and Kenneth J. Arrow, 340. Stanford Mathematical Studies in the Social Sciences. Stanford, Calif.: Stanford University Press.

Benartzi, Shlomo, and Richard H. Thaler. 1995. "Myopic Loss Aversion and the Equity Premium Puzzle." *The Quarterly Journal of Economics* 110 (1) 73–92. https://doi.org/10.2307/2118511

Bianchi, Javier, and Saki Bigio. 2022. "Banks, Liquidity Management, and Monetary Policy." *Econometrica* 90 (1): 391–454. https://doi.org/10.3982/ECTA16599.

Boas, Mary L. 1966. *Mathematical Methods in the Physical Sciences*. New York: John Wiley & Sons, Inc.

Copeland, Adam, Darrell Duffie, and Yilin Yang. 2025. "Reserves Were Not So Ample After All." *The Quarterly Journal of Economics*, 140(1) https://doi.org/10.1093/qje/qjae034

Craig, Ben R., and Sara Millington. 2017. "The Federal Funds Market since the Financial Crisis." *Economic Commentary*, no. 2017-07 (April). https://doi.org/10.26509/frbc-ec-201

707.

Edgeworth, F. Y. 1888. "The Mathematical Theory of Banking." *Journal of the Royal Statistical Society* 51 (1): 113–27. https://www.jstor.org/stable/2979084.

Federal Open Market Committee. 2014. "Policy Normalization Principles and Plans." Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/monetarypolicy/files/FOMC_PolicyNormalization.pdf.

Federal Open Market Committee. 2019. "Statement Regarding Monetary Policy Implementation and Balance Sheet Normalization." Press Release January 30, 2019. https://www.federalreserve.gov/newsevents/pressreleases/monetary20190130c.htm.

Federal Open Market Committee. 2022. "Plans for Reducing the Size of the Federal Reserve's Balance Sheet." Press Release May 04, 2022 Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/newsevents/pressreleases/monetary20220504b.htm.

Federal Reserve Bank of New York, no date. "Effective Fed Funds Rate." https://www.newyorkfed.org/markets/reference-rates/effr accessed February 3, 2025.

Frost, Peter A. 1971. "Banks' Demand for Excess Reserves." *Journal of Political Economy* 79 (4): 805–25. https://doi.org/10.1086/259789.

Gagnon, Joseph E., and Brian Sack. 2019. "Recent Market Turmoil Shows That the Fed Needs a More Resilient Monetary Policy Framework." *Peterson Institute for International Economics Realtime Economic Issues Watch* (blog). September 26, 2019. https://www.piie.com/blogs/realtime-economic-issues-watch/recent-market-turmoil-shows-fed-needs-more-resilient-monetary.

Ghironi, Fabio, and G. Kemal Ozhan. 2020. "Interest Rate Uncertainty as a Policy Tool." Working Paper 27084. National Bureau of Economic Research. https://doi.org/10.3386/w27084.

Goodfriend, Marvin. 2002. "Interest on Reserves and Monetary Policy." *Economic Policy Review* 8 (May): 77–84. https://fedinprint.org/item/fednep/19917.

Hamilton, James D. 1996. "The Daily Market for Federal Funds." *Journal of Political Economy* 104 (1): 26–56. https://doi.org/10.1086/262016.

Hamilton, James D. 2020. "Perspectives on US Monetary Policy Tools and Instruments." In *Strategies for Monetary Policy*, edited by John H. Cochrane and John B. Taylor. Stanford, California: Hoover Institution Press.

Härdle, Wolfgang 1991 *Smoothing Techniques with Implementation in S.* Springer-Verlag, New York.

Ihrig, Jane, Zeynep Senyuz, and Gretchen C. Weinbach. 2020. "The Fed's 'Ample-Reserves' Approach to Implementing Monetary Policy." Finance and Economics Discussion Series. Board of Governors of the Federal Reserve System. https://doi.org/10.17016/FEDS.2020.022.

Keating, Thomas, Francis Martinez, Luke Pettit, Marcelo Rezende, Mary-Frances Styczynski, and Alex Thorp. 2019. "Estimating System Demand for Reserve Balances Using the 2018 Senior Financial Officer Survey." *FEDS Notes* April 09, 2019. DOI: https://doi.org/10.17016/2380-7172.2327

Keister, Todd, Antoine Martin, and James McAndrews. 2008. "Divorcing Money from Monetary Policy." *Economic Policy Review* 14 (2): 41–56. https://www.newyorkfed.org/research/epr/08v14n2/0809keis.html.

Lagos, Ricardo, and Gastòn Navarro. 2023. "Monetary Policy Operations: Theory, Evidence, and Tools for Quantitative Analysis." Working Paper 31370. National Bureau of Economic Research. https://doi.org/10.3386/w31370.

Langowski, Friederike. 2023 "Do bank reserves affect interest rates when reserves are abundant?"

Lopez-Salido, David and Annette Vissing-Jorgensen. 2023. "Reserve Demand, Interest Rate Control, and Quantitative Tightening." *Federal Reserve Board*, January 10.

Lucas, Deborah. 2022. "Some Fiscal Implications of Federal Reserve Balance Sheet Policies, Revisited." Prepared for the Shadow Open Market Committee meeting at Chapman University, June 23, 2022.

Mester, Loretta J. 2024. "The Fed's Ample Reserves Monetary Policy Operating Framework: It Isn't as Simple as It Looks." Remarks for the Panel on The Conduct of Monetary Policy: Evolution from Free Reserves to the Corridor and Floor Systems, Shadow Open Market Committee 50th Anniversary Conference, Hoover Institution, Stanford University.

Poole, William. 1968. "Commercial Bank Reserve Management in a Stochastic Model: Implications for Monetary Policy." *The Journal of Finance* 23 (5): 769–91. https://doi.org/10.1111/j.1540-6261.1968.tb00316.x.

Porteus, Evan L. 2002. *Foundations of Stochastic Inventory Theory.* Stanford, Calif.: Stanford Business Books.

Schnabel, Isabel. 2024. "The Eurosystem's Operational Framework." Speech at the Money Market Contact Group Meeting, Frankfurt Am Main, March 14, 2024. European Central Bank. https://www.ecb.europa.eu/press/key/date/2024/html/ecb.sp240314~8b609de772.en.html.

Schulhofer-Wohl, Sam and Patricia Zobel, 2019. "Transitioning to an Ample Reserves Regime with Lower Reserves." FOMC memo, March 8, 2019. https://www.federalreserve.gov/monetarypolicy/files/FOMC20190308memo02.pdf

Waller, Christopher J. 2022. "Fighting Inflation with Rate Hikes and Balance Sheet Reduction." a Speech at the Economic Forecast Project, February 24, 2022. University of California, Santa Barbara. https://ideas.repec.org//p/fip/fedgsq/93767.html.

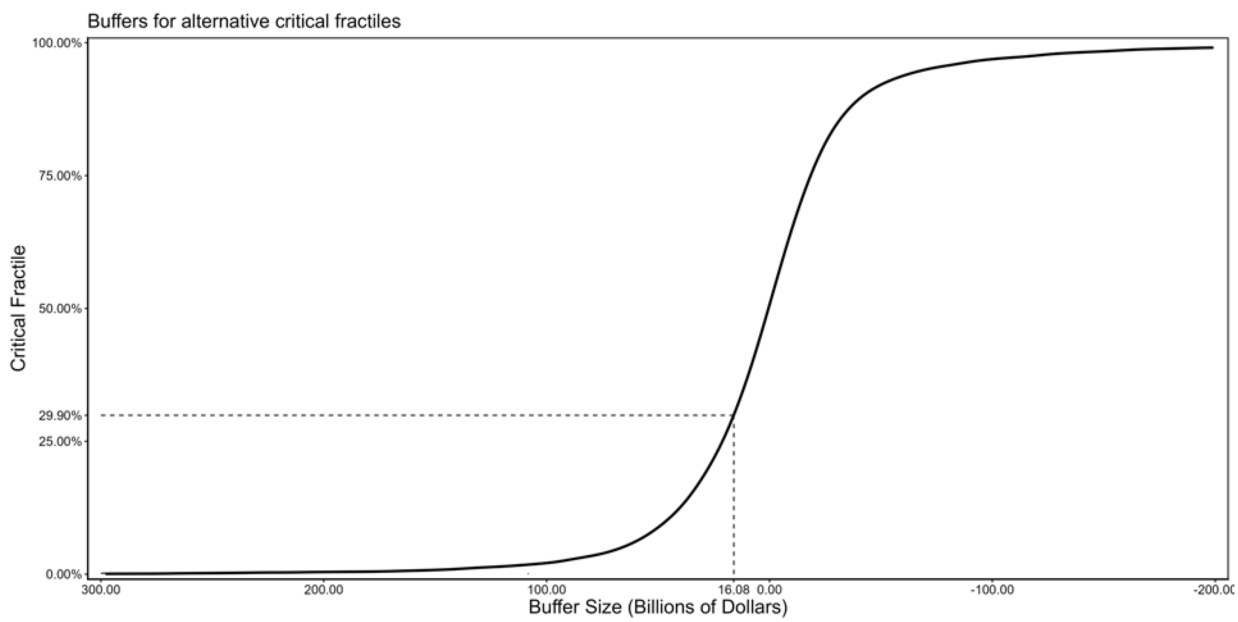

Figure 1: Critical Fractile and Buffer Size. For a given fractile, this shows the optimal buffer, calculated from a kernel density estimate of daily changes in reserves.

6 Appendix:

6.1 Minimizing intervention probability

A plausible class of loss functions, discussed by Afonso et al. (2020), has the social planner minimizing the probability of intervening in the market (preventing the spread from exceeding the ϵ tolerance) plus some cost of having a large balance sheet.

The simple case is where T is high enough that we only have to worry about injecting reserves, the case where increases in supply or decreases in demand, which lower the spread, will only take it to zero and thus stay within the tolerance band. Where is this? We know the kink point in the demand curve is at the point $D - ax = 0$ or $x = \frac{D}{a}$, which implies the target level of reserves must be at $D - aT < \epsilon$ or

$$T \geq \frac{D}{a} - \frac{\epsilon}{a}. \tag{9}$$

or, equivalently, T can be at most $\frac{\epsilon}{a}$ below the kink point $\frac{D}{a}$ of the demand function.

So if condition (9) is satisfied, we only have to worry about supply shocks that decrease reserves and demand shocks that increase demand.

$$Pr[D - a(T + s - \frac{\delta}{a}) > \epsilon]. \tag{10}$$

or

$$Pr[(\frac{D}{a} - T) - \frac{\epsilon}{a} > s - \frac{\delta}{a}]. \tag{11}$$

This has a natural interpretation: the first term ($\frac{D}{a} - T$) is the difference between the kink point of the demand curve and the target rate, and from that is subtracted the reserve equivalent amount of the spread tolerance $\frac{\epsilon}{a}$. This is compared with the reserves equivalent of the supply and demand shocks, which, if negative enough, will force the spread (and thus the interest rate) above the tolerance range.

Given the intervention probability in equation (11), the loss function used by Afonso et al. (2020), which trades off the probability of intervening against the size of the balance sheet, takes the following form:

$$min_T\ L = Pr[(\frac{D}{a} - T) - \frac{\epsilon}{a} > s - \frac{\delta}{a}] + kT. \tag{12}$$

This form assumes that the Fed doesn't care if the rate fluctuates *within* the tolerance range, and that there is a fixed cost of intervening, unrelated to the required size of the intervention. Letting F be the cumulative distribution function (f the density)

$$\frac{\partial L}{\partial T} = \frac{\partial}{\partial T}\{F[(\frac{D}{a} - T) - \frac{\epsilon}{a} > s - \frac{\delta}{a}] + kT\}. \tag{13}$$
$$= -f + k = 0.$$

More compactly,

$$f = k. \tag{14}$$

The second-order conditions are satisfied when $f' > 0$.

The Committee should increase the target level of reserves until the increased probability of intervening just equals the increased cost of a larger balance sheet. The cost of intervening here is normalized to 1, so k, the cost of a larger balance sheet, is in terms of the intervention cost.

A simple comparative static result is also illuminating. Implicitly differentiating (14), we have

$$\frac{dT}{da} = -\frac{Tf'}{af'} = -\frac{T}{a}. \tag{15}$$

This indicates that a steeper demand curve implies that the target balance sheet should be larger.

This loss function is not fully satisfactory, however. The tolerance range, ϵ, is given exogenously, and while in principle it might be determined by the legislature or otherwise specified in advance, ideally it would arise from the costs of trading off interest rate variability, intervention costs, and balance-sheet size. Otherwise, we can have both a small balance sheet and no intervention, provided we put up with a variable funds rate.[8] The inventory approach of Section 3 explicitly considers the trade-off.

6.2 Initial stock and set-up costs

The single-period case has an unrealistic element in that it posits that a per dollar cost of adding to the balance sheet must be applied to the entire balance sheet, whereas intuition suggests it should be closer to an adjustment cost, responding to a shock that is small relative to the overall size of the balance sheet. This can be easily accommodated by defining the initial level of reserves as S, before the FOMC adjusts reserves. Then the objective function becomes

$$G(T,S) = c(T-S) + L(T) = g(T) - cS, \tag{16}$$

where $g(T)$ is defined by

$$g(T) = cT + \int_{-\infty}^{\infty} c_B(T+s)f(s)ds + \int_{-\infty}^{A-T} c_p(A-T-s)f(s)ds. \tag{17}$$

If no action is taken, the cost is $G(S,S)$ and so the cost savings of moving to T is $G(S,S) - G(T,S)$, which should be positive if it is worth moving from S to T. Note that $G(S,S) - G(T,S) = g(S) - g(T)$. But if we take T to be the optimal level T^* it follows from the linearity of marginal costs and the convexity of $g()$ that it is optimal to move from S to T^*.[9] Of course, fixed costs of changing the portfolio will affect this result, leading to a zone of inaction (Porteus, 2002, Chapter 9): if there is a fixed set-up cost c_T of adjusting the balance sheet, the cost savings, $g(S) - g(T)$, must be larger than the fixed cost c_T or no adjustment takes place.

6.3 A discrete-time dynamic inventory approach

Consider a dynamic version of the inventory question, where the balance sheet is carried over to the next period. This follows Chapter 4 of Porteus (2002) so the description will be concise. As before, if the reserves are below the ample level, there is a per unit cost c_p and reserves above the ample level bear a per unit cost c_B. There is a per unit cost of acquiring assets (respectively, inventory) of c. The state variable is x_t, the level of reserves before the central bank intervention, and y_t is the level of reserves after the intervention (but before the shock), so that $x_{t+1} = y_t + s$. There is a terminal value function, where the final reserves of size x are valued at $v_{\mathbb{T}}(x) = -c(x - A)$. The

[8] Ghironi and Ozhan (2020) discuss using the tolerance range as a policy tool itself.

[9] It follows from the convexity of $g(b)$, and since $g'(y) = (c + c_B) - c_p F(y)$ from (5), and $g''(y) = c_p f(y) \geq 0$.

balance sheet must be brought up to the ample level, but anything above that is sold at cost c. This is similar to the assumption of back-orders in the inventory literature. The one-period discount factor is β and shocks are a random variable s with density f and distribution F.

This problem has a recursive formulation and can be solved via backward induction. Let $v_t(x)$ denote the minimum expected cost starting in period t with reserves at level x. The optimality equations become

$$v_t = \min_y \{c(y-x) + L(y) + \beta \int_{-\infty}^{\infty} f_{t+1}(y+s)f(s)ds\}. \tag{18}$$

Letting

$$G_t(y) := cy + L(y) + \beta \int_{-\infty}^{\infty} f_{t+1}(y+s)f(s)ds \tag{19}$$

then the optimality conditions can be rewritten as

$$v_t = \min_y \{G_t(y) - cx\}. \tag{20}$$

Note that if f_{t+1} is convex, $G_t(y)$ is convex, since it is the sum of three convex functions. Hence a y that minimizes $G_t(y)$ gives an optimal level of reserves. Furthermore, f_t is convex, following from convexity preservation under minimization, theorem A.4 of Porteus (2002).

The optimal level can be found more explicity. As a preliminary step, consider the problem of the last period, assuming reserves are at zero. The problem is to minimize

$$cy + L(y) - \beta \int_{-\infty}^{\infty} c(y-s)f(s)ds = c(1-\beta)y + L(y) + \beta c\mu. \tag{21}$$

As above, $L(y) = \int_{-\infty}^{\infty} c_B(y+s)f(s)ds + \int_{-\infty}^{A-y} c_p(A-y-s)f(s)ds$. Then letting $g(b) := c(1-\beta)y + L(y)$ (21) can be written as

$$g(y) + \beta c\mu. \tag{22}$$

Thus, the optimal balance-sheet size S solves

$$g'(S) = 0. \tag{23}$$

Using Leibniz's rule to show that $L'(y) = -c_P + (c_B + c_P)F(y)$ the optimal balance-sheet level is defined implicitly as

$$F(A-S) = \frac{c_B + (1-\beta)c}{c_p}. \tag{24}$$

For this to be finite, the fractile must fall between zero and one, and for that it is sufficient if the cost of a balance sheet below the ample level, c_p, and the cost of holding excess balances, c_B, are both greater than zero, and that the penalty for a low balance sheet is greater than the discounted cost of acquiring and holding assets (otherwise it is optimal to do nothing) or $c_p > (1-\beta)c + c_B$.

So far, equation (24) is only the solution for the last period of the problem. However, we can show that the optimal value functions f_t also have the same slope as the terminal value function, $v_{\mathbb{T}} = -c$, and so (24) will be optimal in each period.

$$f_N(x) = G_N(S) - cx \tag{25}$$

Hence,

$$f'_N(x) = G'_N(S) - c. \tag{26}$$

So $f_N(b)$ has slope $-c$. This argument of course extends back to previous time periods and establishes the recursion.

6.4 A non-negative balance sheet

For the one-period problem, impose the condition that the balance sheet must be non-negative or, equivalently, that the balance-sheet holding cost only applies to a non-negative balance sheet (for example, if a major cost is the interest paid on reserves). The analogue of the basic equation (17) becomes

$$g(T) = cT + \int_{-T}^{\infty} c_B(T+s)f(s)ds + \int_{-\infty}^{A-T} c_p(A-T-s)f(s)ds. \tag{27}$$

Differentiating via Leibniz's rule yields

$$c + c_B[1 - F(-T)] - c_p F(A-T) = 0. \tag{28}$$

This involves two fractiles, so it is not as easily interpretable as the critical fractile (6) but the comparative statics are not difficult to compute. Differentiating implicitly,

$$\frac{dx}{dc_B} = \frac{-[c_B f(-x) + c_p f(A-x)]}{1 - F(-x)}. \tag{29}$$

$$\frac{dx}{dc_p} = \frac{[c_B f(-x) + c_p f(A-x)]}{F(A-x)}. \tag{30}$$

More explicit solutions can be obtained by assuming a functional form for the distribution. For example, letting $F \sim \text{Uniform}[-K, K]$ results in

$$x = \frac{K}{C_p - c_B}[c_p - c_B - 2c] + \frac{Ac_p}{c_p - c_B}. \tag{31}$$