Monetary Policy Implementation in a Negative Rate Environment

Michael Boutros

Duke University

Jonathan Witmer

Bank of Canada

Abstract

To analyze monetary policy implementation in a negative rate environment, we add the option to exchange central bank reserves for cash to the standard workhorse model of monetary policy implementation (Poole, 1968). Importantly, we show that monetary policy can be constrained when the target overnight rate is below the yield on cash. At this point, the overnight rate equals the yield on cash instead of the target rate. Modifications to the implementation framework, such as a reserve requirement that varies with cash withdrawals, can help restore the implementation of monetary policy such that the overnight rate equals the target rate.

Keywords: Interest rates, Monetary policy implementation, Monetary policy framework.

1. Introduction

Central banks have significantly altered their monetary policy implementation frameworks in the aftermath of the financial crisis. First, quantitative easing resulted in an increase in central bank reserves, which significantly changed trading incentives and behavior in the market for overnight reserves. For example, in 2008, the Federal Reserve introduced interest on reserves as a way to maintain influence over the overnight rate because of a significant increase in reserves (Klee et al. (2016)). More recently, several central banks, including the European Central Bank (ECB), the Swiss National Bank (SNB), and the Bank of Japan (BoJ), have adopted negative policy rates.

Central banks implement monetary policy differently, and it is not apparent how differences in monetary policy implementation frameworks matter in a negative rate environment. Most central banks operate by setting a target for the overnight interest rate, along with rates

☆We thank Jason Allen, Greg Bauer, David Cimon, Falko Fecht, Corey Garriott, Todd Keister, Chris Sutherland, Jing Yang, our anonymous referee, and seminar participants at the Bank of Canada, European Central Bank, and Duke University for helpful comments. All remaining errors are our own.

on standing facilities through which participants can borrow from or deposit with the central bank (Borio, 1997). Some central banks, like the Bank of Canada, operate a corridor system whereby the target rate is in the middle of a corridor bounded by the (higher) borrowing rate and the (lower) deposit rate (Bank of Canada, 2015). Others operate a floor system, so named because the target rate is equal to the deposit rate at the bottom of the interest rate corridor. Some central banks have even adapted their frameworks as they lowered their policy rates into negative territory. The Swiss National Bank, for instance, has transitioned to a tiered system for the remuneration of deposits (Swiss National Bank, 2014), where a predetermined level of deposits is exempt from the negative deposit rate and compensated at a rate of zero. Any deposits above this amount are compensated at a negative rate, meaning banks pay the Swiss National Bank for these deposits. How were these changes important for the implementation of monetary policy in a negative rate environment?

Monetary policy implementation is concerned with how short-term (usually overnight) interest rates are determined and is the starting point of the monetary policy transmission mechanism. Understanding the impact of negative rates on monetary policy implementation is of practical importance since negative interest rates are becoming more common. Over USD 13 trillion of sovereign bonds has now traded at negative rates (Whittall and Goldfarb, 2016). And sovereign bond yields in some countries were negative beyond ten years of maturity, suggesting that negative rates were not expected to be a passing phenomenom.

As recently as the end of 2016, central banks with negative rates have a wide range of deposit rates, ranging from -5 bps to -125 bps (Table 1). This could reflect differences in the effective lower bound (ELB) in these countries, which could be related to differences in monetary policy implementation frameworks. It also could reflect the fact that some central banks have not yet hit their ELB, given that yields on cash may be more negative than current overnight rates. This may still be above the negative yield on cash after incorporating the costs of holding and using cash: estimates of the costs of storing and using cash could range from 25 bps (Witmer and Yang, 2016) up to 200 bps (Viñals et al., 2016).

Our paper introduces the ELB to the academic literature on monetary policy implementation by including the option to exchange central bank reserves for cash in a model of monetary policy implementation (e.g., Poole (1968); Bech and Keister (2013)). The opportunity to invest in cash implies an effective lower bound or constraint on overnight interest rates (e.g., Witmer and Yang (2016)), and thus presents a potential obstacle to the implementation of monetary policy. To the best of our knowledge, no other model has considered how the zero lower bound and negative interest rates can impact monetary policy implementation.

Our model contributes several new insights to this literature. First, the central bank target rate must be above the ELB. The marginal cost of lending an extra dollar in the

Table 1: Negative Central Bank Rates as of November 2016 (bps)

Country	Negative Rates Introduced	Lending Rate	Deposit Rate
Danmarks Nationalbank	July 2012*	5	-65
European Central Bank	June 2014	25	-40
Swiss National Bank	Dec. 2014	50	-75
Swedish Riksbank	Feb. 2015	25	-125
Bank of Japan	Jan. 2016	10	-10
Hungarian National Bank	March 2016	115	-5

*The Nationalbank temporarily raised rates into non-negative territory between April and September 2014.

overnight market – the overnight rate – is equal to the respective probabilities of accessing the two central bank standing facilities at the end of the day multiplied by their respective rates (Bindseil, 2001), with these probabilities determined by the uncertainty inherent in the payment shock and the participant's position just prior to the shock. Thus, a participant will convert to cash only if the marginal benefit of converting to cash (i.e., the return on cash) is greater than what they would receive from lending the funds in the overnight market. Thus, the yield on cash does not impact the overnight rate as long as the target for the overnight rate is above the ELB.

Second, if the yield on cash is above the overnight target rate but below the central bank lending rate, the overnight interest rate will equal the yield on cash.[1] Intuitively, participants would prefer to withdraw and invest in higher-yielding cash rather than lend to other participants at the target rate. Their cash withdrawals will lower the overall amount of reserves in the system. In equilibrium, the amount of reserves will adjust until the overnight rate is equal to the return on cash.

Third, our model is the first to examine monetary policy implementation with a tiered deposit rate that allows the tier thresholds to adjust depending on the cash withdrawals of each participant. Whitesell (2006) shows the conditions under which static tiered rates can be used to steer the overnight rate towards the target overnight rate in a positive rate environment.[2] Our model shows how to extend this model to a negative interest rate environment

[1]When the yield on cash is above the central bank lending rate, participants would want to borrow from the central bank and invest in cash, and there is no overnight market. Participants would not want to lend their funds below a rate they receive on investing in cash, and participants would not want to borrow funds above the rate they could attain when borrowing from the central bank.

[2]Different methods have been proposed and utilized for determining the size of the threshold. Whitesell (2006) suggests that the central bank could set the price of the threshold amount and sell this threshold amount for a fee (e.g., 5 bps of the total size of the quota). Holthausen et al. (2008) propose that the central bank could set the quantity of the threshold, and either determine these limits in the same way as they

by adding the option to exchange reserves for cash. We also allow the tiered thresholds to vary with cash withdrawals, and demonstrate why this feature is important for divorcing the overnight rate from the yield on cash. The Bank of Japan and the European Central Bank, for example, have implemented a tiered remuneration of central bank deposits. Our model shows that it is not the tiered remuneration in and of itself that changes incentives to withdraw from the central bank; rather, it is the fact that this tiered remuneration is a function of cash withdrawals that can disincentivize these withdrawals such that the overnight rate once again equals the target rate.

Finally, we develop a model with reserve requirements that are a function of cash withdrawals. This has not been considered in the literature and has not been implemented by any central bank in the negative rate environment. Within our model framework, we show that a varying reserve requirement is more powerful than a varying tiered remuneration in disincentivizing cash withdrawals. This stronger disincentive occurs because, with some probability, banks may not be fully utilizing their exemption threshold in a tiered remuneration framework, so a change in this exemption threshold has a less powerful effect on their incentives. Thus, according to our model, a varying reserve requirement will better help a central bank maintain its influence over the overnight rate when rates are potentially constrained by the lower bound.

Our paper is related to the literature that examines the impact of frictions on monetary policy implementation since the ELB is a friction that can impact the ability of a central bank to influence the overnight interest rate towards its policy rate. Several papers consider how regulation, and in particular the liquidity regulation of banks, will affect monetary policy implementation and the functioning of money markets (Bech and Keister (2013), Banerjee and Mio (2014), Bonner and Eijffinger (2012), Rezende et al. (2016)). Others examine the effect of search frictions, which can generate predictions about volumes and volatility of overnight rates (Bech and Monnet (2013), Afonso and Lagos (2015), Armenter and Lester (2015)). Another related set of papers also considers how segmentation in the overnight market and differential access to central bank facilities can have an impact on monetary policy implementation (Williamson (2015), Bech and Klee (2011), Armenter and Lester (2015), Martin et al. (2013)). Several of these papers show how the introduction of new tools, such as the Federal Reserve's overnight reverse repurchase facility (ORRP) and term deposit facility (TDF) can work to attenuate the effects due to segmentation. Similarly, we show how alterations to the monetary policy implementation framework can attenuate the impact of the ELB on overnight interest rates.

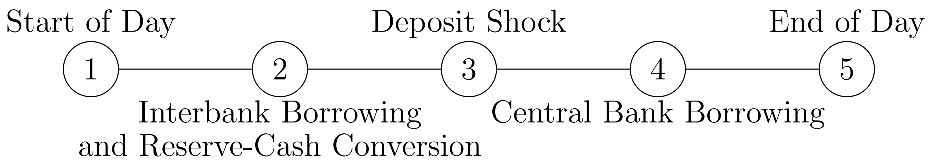
currently determine reserve requirements, or auction the limits to participants.

Since we are examining the ELB, our paper also complements the strand of the literature that analyzes the effect of unconventional tools such as quantitative easing and the resulting large central bank balance sheets and excess reserves on the determination of the overnight interest rate. Goodfriend (2002) and Keister et al. (2008), for instance, point out that when the central bank has excess reserves it essentially has two tools: the interest it pays on reserves, as well as the quantity of those reserves. Kashyap and Stein (2012) suggest that by using both tools distinctly but together, the central bank then has the capability of pursuing two objectives: an inflation objective, and an objective to reduce the externalities created by excessive short-term debt issuance by financial intermediaries. Canzoneri et al. (2017) use a standard DSGE model augmented with a banking sector and interbank trading to explore the optimal interest rate on reserves. Some recent papers discuss how these tools can be used in the exit from unconventional monetary policy (Bech and Klee (2011), Armenter and Lester (2015), Ihrig et al. (2015)). The ELB may limit the ability of the central bank to adjust one of these tools (the interest on reserves), and our model examines how adjustments to implementation frameworks may work to restore this ability.

Given this, we also relate to recent analysis suggesting how the ELB could be lowered or removed. If the central bank restricts conversions of reserves to cash or increases the aggregate stock of paper currency according to a pre-defined rule, a market-determined deposit price of paper currency may develop even if the central bank is still exchanging reserves for cash at par (Goodfriend, 2016). Goodfriend argues that this could, in theory, help to overcome the lower bound on interest rates. However, it may also require changes such that contracts are enforced to be paid in deposits rather than paper currency (Agarwal and Kimball, 2015). Similarly, the central bank could charge a time-varying paper currency deposit fee to eliminate the incentive to withdraw cash to avoid negative interest rates (Agarwal and Kimball, 2015). In our model, an adjustable system of tiered remuneration can also reduce this incentive to withdraw cash. We show how such an adjustable tiered remuneration could be used to reduce the friction associated with the lower bound, at least to a certain degree.

In the next section, we provide the details of our basic model before examining equilibrium impacts in section 3. We consider modifications to our basic model in section 4. In particular, we focus on how a tiered remuneration of central bank reserves (such as is implemented by the Swiss National Bank) can work to mitigate the impact of the ELB on the overnight market. We conclude with a short discussion in section 5.

Figure 1: Model Timing



2. Model

2.1. Banks

As in the static model of Bech and Keister (2013), a continuum of perfectly competitive banks indexed by $i \in [0, 1]$ maximize expected profits. Each day is divided into five stages, as in Figure 1: start of day, interbank borrowing and reserve-cash conversion, realization of deposit shock, central bank borrowing, and end of day. Table 2 shows bank i's balance sheet at each stage of the day, beginning with its balance sheet at the start of the day.

In the next stage, the bank makes two important decisions. Firstly, it participates in the interbank market, and can either become a net borrower from ($\Delta^i > 0$) or net lender ($\Delta^i < 0$) to other banks. Anticipating the uncertain deposit shock in the next period, the bank wishes to manage its reserves such that it minimizes expected costs of accessing the central bank borrowing and deposit facilities in the final period. Secondly, the bank can choose to convert between reserves, R^i, and cash, C^i, via transfers T^i, obeying the constraint that $T^i \geq -C^i$. These two new balance sheet items are reflected in the second panel of Table 2.

Assumption 1. *Every commercial bank's start of day cash balances are equal to zero ($C^i = 0$ for every i) and therefore their daily transfer constraint is simply $T^i \geq 0$.*

> Since profits are linear in cash, this assumption is for ease of exposition and has no effect on commercial banks' level of optimal cash reserves. When the return on reserves is higher than the nominal return on cash, as has been the case for most of history, banks will always choose $T^i = 0$. However, when the return on reserves is sufficiently negative, it may be optimal to convert reserves to cash.

In the next stage of the day, the bank is subjected to a deposit shock ϵ^i that is an independent and identically distributed (i.i.d.) random variable symmetric around zero with CDF G, where $\epsilon^i > 0$ represents a net withdrawal. Drawing on its reserves to accommodate the deposit shock, the commercial bank may now hold too much or too little in reserves relative to the reserve requirement set by the central bank.

In the fourth stage of the day, if the commercial bank's reserves are less than required, the bank borrows X^i directly from the central bank at a rate of r_X. Likewise, any reserves

Table 2: Commercial Bank i's Balance Sheet

Assets	Liabilities
Stage 1: Start of Day	
B^i Bonds	D^i Deposits
C^i Cash	E^i Equity
R^i Reserves	
Stage 2: Interbank Borrowing and Reserve-Cash Conversion	
B^i Bonds	D^i Deposits
$C^i + T^i$ Cash	E^i Equity
$R^i + \Delta^i - T^i$ Reserves	Δ^i Interbank Borrowing
Stage 3: Deposit Shock	
B^i Bonds	$D^i - \epsilon^i$ Deposits
$C^i + T^i$ Cash	E^i Equity
$R^i + \Delta^i - T^i - \epsilon^i$ Reserves	Δ^i Interbank Borrowing
Stage 4 and 5: Central Bank Borrowing (End of Day)	
B^i Bonds	$D^i - \epsilon^i$ Deposits
$C^i + T^i$ Cash	E^i Equity
$R^i + \Delta^i - T^i - \epsilon^i + X^i$ Reserves	Δ^i Interbank Borrowing
	X^i Central Bank Borrowing

Notes: This table itemizes commercial bank i's balance sheet at each stage of the day. Stages four (Central Bank Borrowing) and five (End of Day) are combined. In stage five, bank profits are realized.

in excess of the requirement are deposited with the central bank at the end of the day at a return on excess reserves rate r_R.

Sine most items appear identically on both sides of the balance sheet, each bank's balance sheet identity reduces to a simple expression, relating the asset side (the sum of bonds, B^i, cash, C^i, and reserves, R^i) to the liability side (the sum of deposits, D^i, and equity, E^i):

$$B^i + C^i + R^i = D^i + E^i \tag{1}$$

Assumption 2. *The cost of borrowing from the central bank is always strictly larger than the return on excess reserves:* $r_X > r_R$.

> If this were not the case, then commercial banks could exploit an arbitrage opportunity by borrowing from the central bank at r_X and earning r_R on the borrowed funds by storing them as excess reserves. Poole (1968) shows (under reasonable assumptions) that borrowing from the central bank is essentially always more expensive than borrowing on the interbank market, which motivates commercial banks' interbank activity in stage 2.

Assumption 3. *The return on cash is strictly nonpositive: $r_C \leq 0$.*

This assumption is not critical for the model's results and is made only to analyze the effects of strictly negative interest rates on monetary policy implementation.

Although there is much focus on negative rates below the Zero Lower Bound, there is nothing inherently special about the negativity of negative rates; all that matters is the relationship between the return on cash and the return on reserves. Recent work has estimated the return on cash to be somewhere around -0.5% (Witmer and Yang, 2016), giving rise to the nomenclature Effective Lower Bound.

In our model, the exact levels of the two central bank rates and the return on cash are irrelevant, as what matters is the relationship between them. Both for simplicity and to keep the discussion relevant, we focus on the realistic scenario where the return on cash is zero or slightly negative.

Assumption 4. *Physical cash is not used to settle all deposit shocks.*

Commercial banks do not currently use physical cash to settle most of their large value deposit shocks. We believe this is largely because of the convenience benefits of an electronic payments system. For example, Fedwire has over 9,000 participants and in 2017 processed over 600,000 transfers, on average, each business day.[3] To move a similar amount of cash around would impose very large logistical constraints. Settlement would be delayed due to delivery times, and banks may opt for a delayed net settlement to reduce the volume of transactions. They would have to undertake much more settlement risk if going to delayed net settlement. Storing, transporting, securing, and insuring such a large amount of cash would be possible but would impose a large fixed and ongoing cost. Transportation costs could amount to 1 bps per cash shipment (Witmer and Yang (2016)) which, given the high number of daily transactions, would make using cash uneconomical. Similarly, since it is not structured as a security cash cannot be used as collateral in repurchase agreements.[4]

[3]See `https://www.federalreserve.gov/paymentsystems/fedfunds_ann.htm`. The average daily value of these payments was almost $3 trillion. Further, this is a real-time gross settlement service so settlement is instantaneous.

[4]In a world where interest rates are significantly negative, commercial banks could join together to create a single cash storage and clearing system to make cash payments more economical, but this would still require a large initial investment and may need some sort of central bank oversight if it is a systemic payment system.

2.2. Reserve Requirement

Each bank's total end-of-day reserves must meet the bank's individual reserve requirement, K^i:

$$R^i + X^i + \Delta^i - T^i - \epsilon^i \geq K^i \tag{2}$$

K^i may be set to some constant for all banks, such as zero, constituting an economy-wide reserve requirement. Alternatively, we leave open the possibility that the central bank sets bank-specific conditional reserve requirements. Later, we will show that bank-specific conditional reserve requirements can be used to deter banks from converting reserves to cash in a negative rate environment.

Central bank borrowing X^i must be non-negative and is only utilized if necessary; that is, if total reserves after the deposit shock are less than the reserve requirement. Thus:

$$X^i = \max\{0, K^i - (R^i + \Delta^i - T^i - \epsilon^i)\} \tag{3}$$

Rearranging around the deposit shock, we can determine that central bank borrowing only occurs if:

$$\epsilon^i \geq R^i + \Delta^i - T^i - K^i \equiv \epsilon^i_K \tag{4}$$

where ϵ^i_K is excess reserves for bank i after the interbank market closes (and before central bank borrowing occurs). This equation formalizes the intuition developed earlier: if the deposit shock is greater than the bank's excess reserves following the interbank trading period, the bank will be forced to borrow from the central bank at the end of the day. We can now more succinctly present central bank borrowing as:

$$X^i = \max\{0, \epsilon^i - \epsilon^i_K\} \tag{5}$$

In the appendix, we provide a more general model where cash and reserves can be used to settle payment shocks, and examine the implications when reserves payment shocks disappear. This assumption does not change the main results in the paper.

2.3. Bank Profits

Equation 6 illustrates bank i's profit as a function of the items on its balance sheet.

$$
\begin{aligned}
\pi^i = & \; r_C(C^i + T^i) && r_c \text{ return on cash} \\
& - r_\Delta \Delta^i && r_\Delta \text{ cost of interbank borrowing} \\
& + r_K K^i && r_K \text{ return on required reserves} \\
& - r_X X^i && r_X \text{ cost of central bank borrowing} \\
& + r_R ER^i && r_R \text{ return on excess reserves} \\
& + r_B B^i && r_B \text{ return on bonds} \\
& - r_D(D^i - \epsilon^i) && r_D \text{ cost of deposits}
\end{aligned} \tag{6}
$$

where $ER^i = (R^i + X^i + \Delta^i - T^i - \epsilon^i) - K^i \geq 0$ because of central bank borrowing.

Banks maximize expected profit and so must choose the amount of reserves to convert to cash, T^i, and the level of interbank borrowing, Δ^i, before the deposit shock ϵ^i, is realized. In expectation, bank profits are:

$$\mathrm{E}[\pi^i] = r_B B^i - r_D D^i + r_K K^i - r_\Delta \Delta^i \tag{7a}$$

$$+ r_C(C^i + T^i) \tag{7b}$$

$$+ r_R \epsilon_K^i + (r_R - r_X) \int_{\epsilon_K^i}^{\infty} (\epsilon^i - \epsilon_K^i) \mathrm{dG}(\epsilon^i) \tag{7c}$$

Increasing cash holdings via T^i affects profits directly and indirectly. In (7b), increasing cash holdings directly increases payments of r_C. Under the assumption that $r_C < 0$, a *ceteris paribus* increase in cash decreases profit.

The indirect effect of cash transfers manifests itself in (7c) twice. Firstly, increasing cash decreases excess reserves and their associated returns before central bank borrowing (ϵ_K^i). Secondly, increasing cash increases the likelihood that borrowing from the central bank at rate r_X will be required.

When r_R is greater than the return on cash, increasing cash by reducing excess reserves lowers profits, and $T^i = 0$ is clearly the optimal choice. When r_R is less than the return on cash, however, there may be situations where it is optimal to convert some reserves to cash.

3. Equilibrium

In this section, we formalize the intuition described above into an equilibrium definition, and examine different outcomes under different monetary policy regimes and central bank rates.

Definition. *An equilibrium is a set of individual bank choices* (Δ^i, T^i, X^i) *and interest rate* r_Δ *such that:*

(i) Banks maximize expected profit, (7), subject to their balance sheet constraint, (1), and $T^i \geq 0$.

(ii) The interbank market is a closed system that clears, that is, $\Delta = \int_i \Delta_i \mathrm{d}i = 0$.

3.1. Banks' Maximization Problem

Banks choose T^i and Δ^i to maximize the following Lagrangian:

$$\mathcal{L}^i = \mathrm{E}[\pi^i] + \lambda^i T^i \tag{8}$$

Because the cash transfer constraint may or may not bind in equilibrium (depending on the exogenous rates set by the central bank), we examine the full set of Kuhn-Tucker conditions:

$$r_C - r_R - (r_X - r_R)(1 - G(\epsilon_K^i)) + \lambda^i = 0 \tag{9}$$

$$-r_\Delta + r_R + (r_X - r_R)(1 - G(\epsilon_K^i)) = 0 \tag{10}$$

$$\lambda^i T^i = 0 \tag{11}$$

$$\lambda^i \geq 0, T^i \geq 0 \tag{12}$$

3.2. Aggregation

Equations (9) and (10) show that in equilibrium, all banks choose T^i and Δ^i to yield the same ϵ_K^i and λ^i. Thus regardless of a bank's initial reserves or (potentially bank-specific) reserve requirement, its trading on the interbank market and decision to convert reserves yield the same threshold value after which central bank borrowing must occur. We now refer to this economy-wide threshold as ϵ_K. Optimal trading for a given bank is expressed in equation 13, while equations 14 and 15 integrate over all commercial banks to arrive at economy-wide aggregates.

$$\Delta^i = \epsilon_K^i - R^i + T^i + K^i \tag{13}$$

$$\Delta = \epsilon_K - \int_i (R^i - T^i - K^i)\mathrm{d}i \tag{14}$$

$$= \epsilon_K - R + T + K \tag{15}$$

Given the equilibrium market-clearing condition that the interbank market is a closed clearing system, we have that:

$$\epsilon_K = R - T - K \tag{16}$$

3.3. Equilibrium Interest Rates

Combining equation (16) with the first-order optimization equations (9) and (10) shows that in equilibrium, interest rates are determined based on the aggregate balance sheet statistics, and no one individual bank's choices.

$$r_C = r_R + (r_X - r_R)(1 - G(\epsilon_K)) - \lambda \tag{17}$$

$$r_\Delta = r_R + (r_X - r_R)(1 - G(\epsilon_K)) \tag{18}$$

$$\lambda T = 0 \tag{19}$$

$$\lambda \geq 0, T \geq 0 \tag{20}$$

Solving the system of equations in (17) – (20) yields two distinct cases that we discuss in the following propositions. First, however, it will be useful to compare equilibrium interbank rates in our model with the equilibrium that would arise if converting reserves to cash was forbidden. Poole (1968) showed that the equilibrium interbank rate will fall between the cost of borrowing from the central bank, r_X, and the return on excess reserves, r_R, depending on the level of aggregate reserves in the economy. We call this the Poole interest rate:

$$r_{Poole} \equiv r_R + (r_X - r_R)(1 - G(R - K)) \tag{21}$$

In a so-called corridor system where there are zero excess reserves (i.e., $R = K$), the Poole rate is the midpoint between the cost of central bank borrowing and the return on excess reserves: $r_{Poole} = \frac{r_X + r_R}{2}$. As excess reserves increase (i.e., $R - K$ gets larger), it is less likely that the deposit shock will be large enough to necessitate central bank borrowing, driving down the interbank rate. When these excess reserves become sufficiently large, the likelihood that the deposit shock will be smaller than excess reserves approaches unity, i.e. $G(R - K) \approx 1$, and the interbank rate equals the return on excess reserves. This is often labeled a floor. In summary, depending on the level of excess reserves, the Poole interest rate will fall somewhere between the return on excess reserves and the cost of borrowing from the central bank.

Proposition 1 summarizes the model's prediction of interbank rates in "regular times," that is, when the Poole interest rate is above the return on cash.

Proposition 1. *When the return on cash is less than or equal to the Poole interest rate, then:*

1. *The optimal amount of cash conversions is zero.*

2. *The equilibrium interbank rate is the Poole rate:* $r_\Delta = r_R + (r_X - r_R)(1 - G(R - K)) = r_{Poole}$.

Proof. When the return on cash, r_C, is less than the interbank rate, r_Δ, then from equations (17) and (18) we have that $\lambda > 0$. Then, from equation (19), this implies that $T = 0$, proving the first part of the proposition. When $T = 0$, $\epsilon_K = R - K$ and substituting this into equation (18) yields $r_\Delta = r_{Poole}$. Similarly, according to the first-order conditions, the return on cash can only equal the Poole rate when $\lambda = 0$ and $T = 0$, in which case the Poole rate and interbank rate are equal. □

We call this the "regular times" scenario because, assuming the return on cash is non-positive, this scenario can only arise when the return on excess reserves is non-negative, which has been the case for most of history.[5] Intuitively, converting reserves to cash earns commercial banks the return on cash, but forces them to increase reserves by borrowing on the interbank market at the Poole rate. If the Poole rate is higher than the return on cash, the profit maximizing strategy involves zero cash conversions.

Proposition 2. *When the return on cash is greater than the Poole interest rate but less than the cost of central bank borrowing, then:*

1. *The optimal amount of cash conversions is greater than zero.*

2. *Cash conversions T adjust until the equilibrium interbank rate is equal to the return on cash:* $r_\Delta = r_C$.

Proof. Substituting (16) into equation (18), $r_\Delta > r_{Poole}$ only when T is greater than zero. When T is greater than zero, it must be the case that $\lambda = 0$, which implies $r_\Delta = r_C$ from equations (17) and (18). □

Intuitively, when the return on cash is greater than the Poole interest rate, it is profitable to convert reserves to cash and make up the shortfall by borrowing on the interbank market at the Poole rate. This is an essentially an arbitrage opportunity that the commercial bank will exploit so long as it remains profitable to do so. For every unit of reserves converted to cash, the marginal profit of the arbitrage opportunity decreases because the interbank rate increases, until the marginal profit is zero and the arbitrage opportunity is no longer profitable.

This is because in the presence of cash transfers, the interbank rate, $r_\Delta = r_R + (r_X - r_R)(1 - G(R - K - T))$, is similar to the Poole rate in form but increases in cash transfers

[5] We emphasize our use of the Poole rate in proposition 1. It is completely possible for the Poole rate to be positive and larger than the return on cash even if the return on excess reserves is negative.

T because the likelihood of central bank borrowing increases. In response, the interbank market rate increases towards the return on cash. Cash transfers T will increase until the cost of replacing converted reserves via interbank loans is exactly equal to the return on cash. Past this point, cash transfers yield a return less than their cost. Thus, in equilibrium the level of cash transfers T is such that r_Δ is equal to r_C.

It is important to note that in this framework, when the Poole rate is below the return on cash, the central bank loses its control of the interbank rate and, to a certain extent, the level of excess reserves. The interbank rate will be wholly determined by the return on cash, a rate over which the central bank generally has no control. Excess reserves are now $R - T - K$, and although the central bank can still influence the level of excess reserves by changing the level of required reserves, K, cash transfers T are wholly determined by the profit-maximizing commercial banks.

3.3.1. Equilibrium Under Different Monetary Policy Regimes

In the previous section we analyzed equilibrium outcomes when the return on cash was above or below the Poole rate. In this section we analyze how the Poole rate changes when the level of aggregate excess reserves, $R - K$, changes. A central bank may explicitly target a specific monetary policy framework, such as setting $R = K = 0$ and using a "corridor system." A central bank may also prioritize other objectives over strictly controlling the level of excess reserves; for example, if a central bank operates a Large-Scale Asset Purchase Program funded by reserves, and does not correspondingly change the level of required reserves, then the level of excess reserves will be determined by the size of the asset purchase program.

Figure 2 illustrates how the aggregate level of reserves interacts with the Poole rate. The curved line is the Poole rate, which at $R = K = 0$ is exactly the midpoint between the cost of central bank borrowing and the return on excess reserves.

The three coloured regions represent the type of equilibrium that occurs for a given set of interest rates (r_X, r_R), return on cash (r_C), and the level of aggregate reserves. The region in which the horizontal return on cash line intersects the vertical aggregate reserves line determines the type of equilibrium.

1. Blue region: the return on cash is higher than the cost of borrowing from the central bank. In this case, the most profitable strategy for a bank would be to borrow from the central bank and hold that amount as cash.[6]

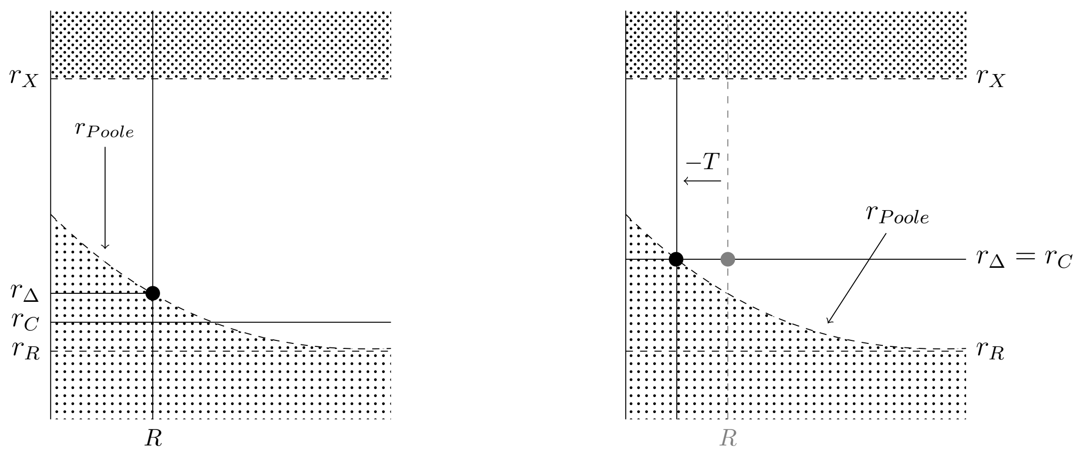
[6] We do not model this scenario explicitly (instead, central bank borrowing is always the minimum required to meet reserve requirements) since it is highly unlikely to occur and the implications are obvious.

2. Green region: the return on cash is higher than the "Poole" rate (but lower than the cost of borrowing from the central bank), in which case some cash holdings are optimal.

3. Red region: the return on cash is lower than the overnight rate that would exist in the absence of cash transfers (the "Poole" rate), which we've already shown will imply zero cash holdings.

The figure's colours are chosen from the perspective of a commercial bank making a decision about converting reserves to cash; simply put, the red region tells the commercial bank not to convert any reserves to cash while the green region says that some cash holdings are optimal.

Figure 2: General Illustration of Equilibrium Regions



(a) Equilibrium with $T = 0$

(b) Equilibrium with $T > 0$

Notes: Panel (a) illustrates the case where the return on cash, r_C, intersects the aggregate level of reserves in the red region (dark circle), indicating to the commercial banks that zero cash conversion is optimal. In these charts, $K = 0$. In panel (b), the return on cash line intersects the aggregate level of reserves in the green region (light circle), signaling to commercial banks that some positive cash holding is optimal. As a result, in equilibrium, T cash conversions from reserves decrease the level of excess reserves) until the return on cash intersects the new $R-T$ line in the red region (dark circle), signaling to commercial banks that zero additional cash conversion is optimal. The equilibrium interbank rate is equal to the return on cash.

Figure 2 illustrates the relationship between the aggregate level of excess reserves, the rates set by the central bank, and the return on cash. The horizontal axis represents the aggregate level of excess reserves, determined by R (we set $K = 0$ to simplify the illustration). If $K > 0$, then the vertical would represent the level of excess reserves, $R - K$. On the far left-hand side, $R = 0$, indicating a corridor system. In this case, the overnight rate that would exist if cash transfers were not possible, r_{Poole}, is the midpoint of the central bank borrowing (r_X) and deposit (r_R) rates. As R increases, we approach a floor system where the endogenous Poole rate is equal to the deposit rate.

If the return on cash is in the red area below r_{Poole}, zero cash conversions are optimal (case three). This is the case illustrated in panel (a), where the aggregate level of reserves, R, is a vertical line and the return on cash, r_C, is a horizontal line. The coloured region where these two lines intersect dictates the equilibrium case that will arise for this combination of monetary policy framework and return on cash. The two lines intersect in the red region, indicating it is optimal for banks to convert zero reserves to cash. Note, however, that with the exact same set of rates but with a higher level of excess reserves, it would be optimal to convert some cash to reserves.

Case two occurs when the return on cash is in the green area between the r_X and r_{Poole} line. This is the case illustrated in panel (b), where the return on cash line intersects the aggregate reserves line, R, in the green region, indicating the economy is out of equilibrium. Commercial banks can earn positive profit by converting cash to reserves, earning $r_C > r_\Delta$, and borrowing on the interbank market to meet reserve requirements. Commercial banks continue to convert cash to returns until equilibrium is reached where cash transfers T are such that the interbank rate is equal to the return on cash. The solid vertical line is the "effective" monetary policy framework $R - T$, illustrating the concept that the cash transfers increase the interbank market rate relative to what it would be in the absence of cash transfers.

4. Modelling Alternative Central Bank Frameworks

In negative rate environments in our model, cash conversions affect the overnight interbank rate and drive it to equal the return on cash. This may be undesirable from the central bank's perspective since the overnight interbank rate is often a key policy target. In order to maintain control of the interbank rate, the central bank can attempt to disincentivize commercial banks from converting reserves to cash. In this section, we model alternative central bank frameworks, some of which have already been implemented, and analyze the equilibrium interbank rate.

4.1. Tiered Remuneration of Central Bank Reserves

In January 2016, the Bank of Japan followed the Swiss National bank by announcing "a three-tier system in which the outstanding balance of each financial institution's current account at the Bank will be divided into three tiers, to each of which a positive interest rate, a zero interest rate, or a negative interest rate will be applied, respectively." In particular, the Bank stated that "... if a financial institution increases its cash holdings significantly, the bank will deduct an increase in its cash holdings from the zero interest-rate tiers of current account balance. Thus, a negative interest rate will be charged on the increase in its cash holdings."

The central bank can regain monetary policy efficacy, even with cash transfers, by using the tiered system to reduce an amount of reserves exempted from being compensated at the lower deposit rate by exactly the same amount as cash conversions. Recall that the exempted reserves can be bank-specific, and in reality this is completely plausible since the central bank maintains an actual account for each commercial bank.

Tiered reserve rates in negative rate environments have two components. The first is an exemption on paying negative rates on deposits up to a certain threshold, beyond which deposits are compensated at the deposit rate. When the deposit rate is negative, compensation at the deposit rate is a cost for commercial banks instead of a source of revenue. The second component is an exemption threshold chosen by the central bank. We consider the effects of both a fixed threshold or a threshold that varies with cash conversions.

4.1.1. Tiered Remuneration with Fixed Threshold

We modify the framework developed in section 2 to add tiered remuneration with a fixed exemption threshold. In particular, commercial banks are compensated at r_M on the first M dollars of reserves they hold at the central bank. Beyond the threshold M, reserves are compensated at the deposit rate, r_R. For simplicity we set the amount of required reserves to zero (i.e., $K = 0$).

Assumption 5. *The exempted rate is between the central bank lending and deposit rates:* $r_R < r_M \leq r_X$.

> This means that to be an exemption, the exempted rate, r_M, should be higher than the central bank deposit rate. If the exempted rate were above the central bank lending rate, banks' willingness to trade will depend on their initial level of reserves and banks will not necessarily choose their interbank activity depending on the same critical values. Banks with positive reserves below the threshold and banks in need of reserves would not have incentive to trade with each other, as they could each receive a better rate by transacting with the central bank.
>
> In a negative interest rate environment, central banks have typically set $r_M = 0$, effectively restricting $r_X \geq 0$. In practice, central banks have not set their standing facility lending rates below zero in a tiered remuneration system.

Bank profits are still described by equation 6 except that excess reserves up to M earn r_M and all reserves past this threshold earn r_R. In expectation, bank profits in this new framework are:

$$\mathrm{E}[\pi^i] = r_B B^i - r_D D^i - r_{\Delta,M}\Delta^i \tag{22a}$$
$$+ r_C(C^i + T^i) \tag{22b}$$
$$+ r_M \int_{-\infty}^{\epsilon_K^i} (\epsilon_K^i - \epsilon^i)\mathrm{dG}(\epsilon^i) \tag{22c}$$
$$+ (r_R - r_M) \int_{-\infty}^{\epsilon_K^i - M^i} (\epsilon_K^i - M^i - \epsilon^i)\mathrm{dG}(\epsilon^i) \tag{22d}$$
$$- r_X \int_{\epsilon_K^i}^{\infty} (\epsilon^i - \epsilon_K^i)\mathrm{dG}(\epsilon^i) \tag{22e}$$

Since $K = 0$, we now have that $\epsilon_K = R - T$. Maximization and aggregation are similar to the model above. The resulting first-order and Kuhn-Tucker conditions are analyzed together:

$$r_{\Delta,M} = r_R(G(\epsilon_K - M)) + r_X(1 - G(\epsilon_K)) + r_M(G(\epsilon_K) - G(\epsilon_K - M)) \tag{23}$$
$$r_C = r_{\Delta,M} - \lambda_M \tag{24}$$
$$\lambda_M T = 0 \tag{25}$$
$$\lambda_M \geq 0, T \geq 0 \tag{26}$$

Given the same definition of equilibrium, the interbank rate is:

$$r_{\Delta,M} = r_R(G(R - T - M)) + r_X(1 - G(R - T)) + r_M(G(R - T) - G(R - T - M)) \tag{27}$$

Once again, it is useful to define the interbank rate that would exist if there were no ability to convert reserves to cash:

$$r_{Poole,M} \equiv r_R(G(R - M)) + r_X(1 - G(R)) + r_M(G(R) - G(R - M)) \tag{28}$$

A special case arises when $r_M = \frac{r_X + r_R}{2}$ and $M = 2R$. By choosing an exemption threshold equal to double the level of aggregate reserves and setting the return on deposits up to the threshold equal to the midpoint of the central bank lending and borrowing rates, $r_{Poole,M} = r_M$. Outside of this example, the tiered Poole rate will be bounded by the central bank lending and borrowing rates. However, conversions of reserves to cash could cause the interbank rate to deviate from this tiered interbank rate. The general cases are summarized in the the next two propositions.

Proposition 3. *In a tiered remuneration system with a fixed exemption threshold, if* $r_{Poole,M} \geq$

r_C, then:

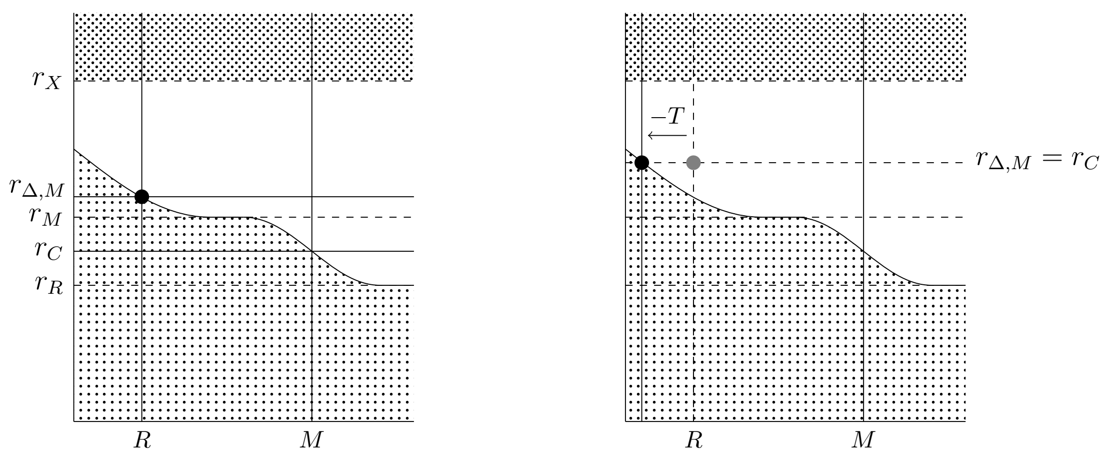
1. *The optimal amount of cash conversions is zero.*

2. *The equilibrium interbank rate is the tiered Poole rate: $r_{\Delta,M} = r_{Poole,M}$.*

Proof. When the interbank rate ($r_{\Delta,M}$) is greater than the return on cash, r_c, then from equation (23), it must be that $\lambda_M > 0$. Then, cash transfers must be zero from equation (25). When $T = 0$, $\epsilon_K = R$, and substituting this into equation (23) yields $r_{\Delta,M} = r_{Poole,M}$. Similarly, according to the first-order conditions, in this case the tiered Poole rate and interbank rate are equal. □

This mirrors the earlier results for a standard operating framework: the interbank market continues to function properly so long as the return on cash is below the interbank rate that would exist in the absence of cash transfers.

Figure 3: Illustration of Equilibrium Regions for Tiered Remuneration



(a) Equilibrium with $T = 0$ (b) Equilibrium with $T > 0$

Notes: Panel (a) illustrates the case where the return on cash, r_C, intersects the level of required reserves, R, in the red region, indicating to the commercial banks that zero cash conversion is optimal. In this case, monetary policy implementation works as normal. In panel (b), the return on cash line intersects the level of required reserves in the green region, signaling to commercial banks that some positive cash holding is optimal. As a result, in equilibrium, T cash conversions from reserves decrease the level of required reserves (i.e., the monetary policy framework) until the return on cash intersects the new $R-T$ line in the red region, signaling to commercial banks that zero additional cash conversion is optimal. The equilibrium interbank rate is equal to the return on cash and monetary policy implementation is constrained. The exemption threshold, M, is unaffected by the cash conversion and remains fixed.

The more interesting cases arises when considering a tiered Poole rate below the return on cash.

Proposition 4. *In a tiered remuneration system with no adjustments to the exemption threshold, when the return on cash is greater than the tiered Poole rate but less than the cost of central bank borrowing, then:*

1. *The optimal amount of cash conversions is greater than zero.*
2. *Cash conversions T adjust until the equilibrium interbank rate is equal to the return on cash: $r_{\Delta,M} = r_C$.*

Proof. Substituting $\epsilon_K = R - T$ into equation (23), $r_{\Delta,M} > r_{Poole,M}$ only when T is greater than zero. When T is greater than zero, it must be the case that $\lambda_M = 0$, which implies $r_{\Delta,M} = r_C$ from equation (24). □

These propositions illustrate that tiered remuneration in and of itself cannot insulate the interbank market from the effect of converting reserves to cash. Interbank rates are effectively floored at the return on cash in both the standard set-up and the tiered remuneration framework.

Figure 3 illustrates the regions where cash conversion will occur in this framework. From the perspective of a commercial bank, it is not optimal to make cash conversions when the aggregate level of reserves intersects the horizontal line defining the return on cash is in the red region, as illustrated in panel (a).

Panel (b) of figure 3 shows the case where it is optimal for commercial banks to convert reserves to cash. They will convert reserves until the interbank rate is equal to the return on cash, which occurs exactly when the return on cash intersects the monetary policy framework in the red region. The mechanism is similar to the mechanism described in the earlier section for the standard implementation framework.

4.1.2. Tiered Remuneration with Varying Threshold

Consider now the same remuneration framework except that the central bank adjusts the exemption threshold used to calculate the tiered remuneration. Specifically, each bank's individual exemption threshold, M^i, varies with that bank's cash conversions:

$$M^i = \bar{M} - aT^i \tag{29}$$

We have generalized the penalty for cash conversions such that the exemption threshold can be lowered by an amount greater than the amount of reserves converted to cash to create a greater disincentive. Central banks that have implemented a varying threshold framework lower this threshold one-for-one (i.e., $a = 1$).

Recall that ϵ_K^i is the threshold relative to ϵ^i that determines whether or not central bank borrowing is required. With M^i defined as above:

$$\begin{aligned}\epsilon_K^i - M^i &\equiv R^i + \Delta^i - T^i - (\bar{M} - aT^i) \\ &= R^i + \Delta^i - (1-a)T^i - \bar{M}\end{aligned}$$

Now, cash conversions are less appealing because they lower the stock of reserves which are compensated at the relatively higher remuneration rate. However, converting to cash doesn't lower the threshold after which central bank borrowing is required, so there still may be some incentive to convert to cash, especially if banks will not be fully depositing reserves up to the threshold with certainty. Inserting the new threshold into individual banks' profit function in equation (22) yields a slightly modified expected profit function:

$$\mathrm{E}[\pi^i] = r_B B^i - r_D D^i - r_{\Delta,M}\Delta^i \tag{30a}$$
$$+ r_C(C^i + T^i) \tag{30b}$$
$$+ r_M \int_{-\infty}^{R^i+\Delta_i-T^i} (R^i + \Delta_i - T^i - \epsilon^i)\mathrm{dG}(\epsilon^i) \tag{30c}$$
$$+ (r_R - r_M) \int_{-\infty}^{R^i+\Delta_i-(1-a)T^i-\bar{M}} (R^i + \Delta_i - (1-a)T^i - \bar{M} - \epsilon^i)\mathrm{dG}(\epsilon^i) \tag{30d}$$
$$- r_X \int_{R^i+\Delta_i-T^i}^{\infty} (\epsilon^i - (R^i + \Delta_i - T^i))\mathrm{dG}(\epsilon^i) \tag{30e}$$

Maximizing expected profits yields two first-order conditions. Combining them, aggregating across all commercial banks, and using the fact that $\Delta = 0$, the optimality conditions can be expressed as:

$$r_{\Delta,M} = r_R(G(R-(1-a)T-\bar{M})) + r_X(1-G(R-T)) + r_M(G(R-T) - G(R-(1-a)T-\bar{M}))$$
$$r_C = r_{\Delta,M} + a(r_M - r_R)G(R-(1-a)T-\bar{M}) - \lambda_M$$

As expected, this framework continues to yield that monetary policy implementation implementation is unconstrained when central bank rates are positive.

Proposition 5. *In a tiered remuneration system with a dynamically adjusting exemption threshold, when* $r_C < r_{Poole,M} + a(r_M - r_R)G(R - \bar{M})$*, then:*

1. *The optimal amount of cash conversions is zero.*

2. *The equilibrium interbank rate is the tiered Poole rate:* $r_{\Delta,M} = r_{Poole,M}$*.*

Proof. When the interbank rate ($r_{\Delta,M}$) is greater than ($r_C + a(r_R - r_M)G(R - \bar{M})$), from the first first-order condition λ_M must be greater than 0, and this implies that cash transfers must be zero. When $T = 0$, according to the first-order conditions in this case the tiered Poole rate and interbank rate are equal. □

The intuition for this result can be seen in the second optimality condition, which describes the marginal benefit and marginal cost of converting a dollar of reserves to cash. With this conversion, they will gain the return on cash (on the left hand side of the equation). However, the marginal cost contains two terms. First, they would lose what they could otherwise have earned by trading the dollar of reserves to another bank, $r_{\Delta,M}$. The second term is positive and represents the effect of reducing the exemption threshold for converting a dollar of reserves to cash. The exemption threshold changes by a dollars, and they will earn a lower r_R (instead of r_M) on the a dollars the exemption threshold changes by. Since, depending on the amount of reserves in the system, they may not be fully utilizing their exemption threshold, this penalty is multiplied by the probability that they are using the exemption threshold.

The marginal cost of converting reserves to cash, according to the second condition for cash transfers, is simplified when $a = 1$:

$$r_C = r_M G(R - T) + r_X(1 - G(R - T)) + \lambda$$

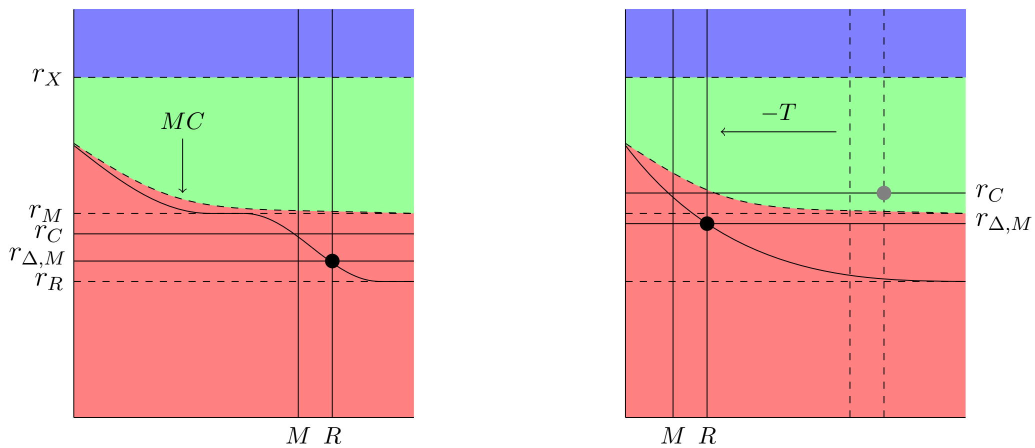
There will be no cash transfers if the return on cash is less than this marginal cost of converting reserves to cash (when T is set to 0). Thus, as long as this weighted average of r_M and r_X is higher than the return on cash, banks do not have an incentive to convert reserves to cash. Otherwise, banks will convert T excess reserves to cash until the marginal benefit of cash conversions equals the marginal cost.

In a system of tiered remuneration with abundant reserves, the maximum amount the interbank rate can be below the return on cash is equal to $r_M - r_R$ (the difference between the two lines at the far right hand side of Figure 4). Since reserves are abundant, $G(R) \approx 1$ and the cash transfer first order condition reduces to $r_C \leq r_M$. Similarly, the first order condition for interbank trading reduces to $r_\Delta = r_R$. Subtracting the second first order condition from the first and rearranging yields: $r_C - r_\Delta \leq r_M - r_R$.

The Swiss National Bank operates a framework of this nature. Exempted reserves below the threshold earn 0 bps (i.e., $r_M = 0$ bps) and reserves held above the threshold earn -75 bps (i.e., $r_X = -75$ bps). Thus in the Swiss case this suggests the interbank rate could trade up to 75 bps below the return on cash. If a central bank with abundant reserves desired a lower interbank rate, it would have to either increase r_M to the top of its corridor, increase the width of its corridor, or increase the penalty a for cash transfers.

Once r_M is near the top of the corridor, or if it is initially set near the top of the corridor, then our model shows why increasing it may not provide added flexibility. Increasing the width of the corridor or increasing the penalty a may have political economy implications that are outside the scope of our model but are still of first-order importance. For example, increasing the width of the corridor makes the borrowing rate from the central bank more penal relative to the interbank rate. Though this will not affect the technical implementation of monetary policy, it may be politically untenable to the banking industry.

Figure 4: Illustration of Equilibrium Regions with Tiered Remuneration and Varying Threshold



(a) Equilibrium with $T = 0$ (b) Equilibrium with $T > 0$

Notes: In this figure, the dashed curved line (MC) represents the marginal cost of converting an additional dollar of reserves to cash, and the difference between this line and the demand curve for interbank reserves represents the additional flexibility gained by varying exempted reserves with cash demand. Panel (a) illustrates the case where the return on cash, r_C, intersects the level of required reserves, R, in the red region, indicating to the commercial banks that zero cash conversion is optimal. In this case, monetary policy implementation works as normal. In panel (b), the return on cash line intersects the level of required reserves in the green region, signaling to commercial banks that some positive cash holding is optimal. As a result, in equilibrium, T cash conversions from reserves decrease both reserves and the exemption threshold until the return on cash intersects the level of reserves in the red region, signaling to commercial banks that zero additional cash conversion is optimal.

In Figure 4 we add a line to illustrate the marginal cost of converting an additional dollar of reserves to cash. Reserves will not be converted to cash if there are enough reserves such that this cost is greater than the return on cash. Thus, the area where cash conversion does not occur can be redefined based on the return on cash relative to this line representing the marginal cost of converting reserves to cash.

The difference between this line defining the marginal cost of converting reserves to cash and the demand curve for interbank reserves represents the additional flexibility that a central bank can gain by varying exempted reserves with cash demand.[7]

The difference between the cost of converting to cash and the interbank rate is the maximal amount that the interbank rate can be below the return on cash. In the figure, this is the difference between the line labelled MC and the line representing the interbank rate. To see this, when the threshold moves one-for-one with cash transfers (i.e., $a = 1$), the first order conditions for cash transfers suggests $r_C \leq r_M G(R) + r_X(1 - G(R))$ before this constraint binds.

Proposition 6. *In a tiered remuneration system with a dynamically adjusted exemption threshold, when the return on cash is less than the cost of central bank borrowing and* $r_C > r_{Poole,M} + a(r_M - r_R)G(R - \bar{M})$*, then:*

1. *The optimal amount of cash conversions is greater than zero.*

2. *Cash conversions* T *adjust until:* $r_C = r_{\Delta,M} + a(r_M - r_R)G(R - \bar{M})$*.*

Proof. This follows directly from the second first-order condition. □

In this second equilibrium, cash conversions occur but do not cause the interbank rate to equal the return on cash. Instead, the interbank rate approaches a rate that is less than the return on cash. The higher r_M, the return paid on exempted reserves, the less cash conversions push the interbank rate towards the return on cash.

It is possible that cash transfers become large enough such that the redemption threshold, $\bar{M} - T^i$, becomes negative. In our analysis, we assume that the central bank is able to implement the equivalent of this redemption threshold when reserves are negative. This would in effect be assessing a penalty on net cash transfers. If the exemption threshold is restricted to be positive, then banks would have an incentive to convert reserves to cash if the benefit of doing so exceeds the lost benefit from utilizing the exemption threshold. This is more likely to be the case if the exemption threshold is small.

4.1.3. Optimality of a

A central bank should want an interbank rate equal to its target for that rate that is consistent with its monetary policy objectives. In our model, this can be achieved by setting $a = \infty$, which effectively bans cash conversions. In this case, the return on cash has no impact on monetary policy implementation and there is no lower bound. To see this, from the first order conditions in the model, there are no cash withdrawals and the interbank rate is unconstrained when $r_R G(R - \bar{M}) + r_X(1 - G(R)) + r_M[G(R) - G(R - \bar{M})] >=$

[7] The figure illustrates the case when $a = 1$. This marginal cost line is even higher if a is increased.

$r_c + a(r_R - r_M)G(R - \bar{M})$. When $a = \infty$, the right hand side of this inequality is equal to $-\infty$, suggesting no constraint.

However, Agarwal and Kimball (2015) provide several reasons why restricting the quantity of cash withdrawn may be undesirable. Restrictions are intended to prevent cash withdrawals for storage purposes but could also have the unintended consequence of restricting withdrawals for spending. Quantity restrictions on withdrawals could also result in an effective price of paper currency (relative to reserves) that is different from 1 and that moves abruptly. It could also give people an incentive to hoard paper currency, especially if they foresee the introduction of, or an increase in, the penalty for withdrawing. This could impact the velocity of money which may be unintended for the central bank.

To put this in the context of our model, we assume that a central bank will choose a value of a to minimize the squared difference between the overnight rate and the Poole rate (i.e., the rate that would exist if cash transfers were not possible). Specifically, the central bank would want to choose a value of a to minimize:

$$(r_{\Delta,M} - r_{Poole,M})^2$$

Effectively, the first order conditions from this minimization boils down to choosing a such that no transfers occur, $T = 0$. Put another way, this difference is minimized when $r_{\Delta,M} = r_{Poole,M}$, which by definition occurs when $T = 0$. From the Lagrangian in the first order conditions, $T = 0$ when $\lambda \geq 0$. Combining the two first order conditions when T and λ take on these values and rearranging yields:

$$r_c \leq r_R G(R - \bar{M}) + r_X(1 - G(R)) + r_M[G(R) - G(R - \bar{M})] + a(r_M - r_R)G(R - \bar{M})$$

The term on the left hand side of the equation is the benefit from converting an extra dollar of reserves to cash. The first three terms on the right hand side represent the cost to converting reserves to cash when the reserves threshold does not change. This is what the interbank rate would be in the absence of the option to convert reserves to cash (i.e., $r_{Poole,M}$). The last term on the right hand side is the penalty the bank receives when the exemption threshold varies with cash conversions. For every dollar converted to cash, the exemption threshold changes by a dollars. Instead of earning the higher exemption rate on these a dollars, the bank earns the lower deposit rate. This only affects the bank so long as its reserves are above the exemption threshold. When reserves are below the threshold, changing the threshold does not change the banks compensation. Therefore, this final term includes the probability that the bank's reserves are above the threshold, $G(R - \bar{M})$. Rearranging this equation to solve for a produces:

$$a \geq \frac{r_c - r_{Poole,M}}{(r_M - r_R)G(R - \bar{M})}$$

The central banks that have implemented a varying reserve framework have chosen $a = 1$ in terms of penalizing withdrawals. This may be out of simplicity, but it also is consistent with the optimal value in the model. The central banks that have implemented tiered remuneration with varying thresholds have abundant reserves. In an environment of abundant reserves, $r_{Poole,M} = r_R$ and $G(R - \bar{M}) = 1$. In this case, the solution for a simplifies to: $a \geq \frac{r_c - r_R}{r_M - r_R}$. Both the return on cash as well as the rate on exempted returns are zero, or close to it. In this case, the equation further simplifies to $a \geq 1$. Inside the model, any value of a greater than one could achieve the desired target. Outside the model, however, there are important political economy effects of choosing a higher value of a, which may speak to why central banks with this framework have chosen $a = 1$.

4.1.4. Required Reserves with Varying Threshold

Instead of tiered remuneration of central bank reserves, a central bank could implement a system of required reserves, where the reserve requirement, K^i, adjusts depending on the amount of cash withdrawals:

$$K^i = \bar{K} - T^i \tag{31}$$

Proposition 7. *In a required reserve system with a dynamically adjusting reserve requirement, when $r_K \geq r_C$, then:*

1. *The optimal amount of cash conversions is zero.*
2. *The equilibrium interbank rate is equal to the Poole rate: $r_\Delta = r_{Poole} = r_R + (r_X - r_R)[1 - G(R - \bar{K})]$*

Proof. See Appendix A. □

Under this system, excess reserves never change because converting one unit of reserves to cash lowers both total and required reserves by one unit. The commercial bank earns an additional r_C and loses r_K on the newly converted unit of cash. From this, it is clear to see why an equilibrium of zero cash transfers exists whenever $r_K \geq r_C$.

The primary advantage of such a system is that it can produce even tighter control over the overnight rate given that the overnight rate is the same as it would be when interest rates are above the return on cash (i.e., the Poole rate). Propositions 4 and 6 gave conditions under which the central bank would lose control of the interbank rate in fixed and varying tiered remuneration systems. No such conditions exist for this system. As long as the return on required reserves is greater than or equal to the return on cash, the interbank market will clear at the Poole rate.

We emphasize that the choice of return on required reserves is independent of the central bank's decisions of the return on (excess) reserves, r_R, and the cost of central bank borrowing, r_X. Even if the central bank is operating in negative territory, under this system, all that's required is the return on required reserves be higher than the return on cash. Since the return on cash is less than or equal to zero, setting the return on required reserves greater than zero yields the desired result, even if both the return on excess reserves and the cost of central bank borrowing are both negative. In both the varying reserve requirement and varying exemption threshold frameworks, however, as rates become more negative, they'll likely have an increasingly negative impact on central bank profits. Thus, although the model shows that a central bank is capable of achieving its target rate with a varying reserve requirement, the cost to the central bank of this capability increases as rates become deeply negative.

This system provides the additional advantage that it should be relatively easy to implement, since many central banks already have required reserves. Because of the relative ease of implementation and more powerful disincentivization of cash withdrawals, a varying reserve requirement may be more advantageous than tiered remuneration for continuing to implement monetary policy in a negative rate environment.

5. Conclusion

Our model illustrates how tiered remuneration of central bank deposits, in and of itself, does not relax the lower bound constraint much. However, it is the ability to vary the amount of reserves exempted from being compensated at the lower deposit rate with the amount of reserves converted to cash that allows the interbank market to clear at a rate below the return on cash. The model shows how using a reserve requirement that varies with conversions of reserves to cash can allow the interbank rate to clear below the return on cash in a larger number of scenarios.

Implementing the tiered remuneration system proposed in this paper appears to be feasible. Some central banks have implemented tiered remuneration, suggesting that upfront transition costs (e.g., updating IT systems) are manageable. The Swiss National Bank, for example, implemented a form of tiered remuneration and policy rates have traded close to their target during and since the transition. The system works despite heterogeneity in participants and in how the rules are applied.[8] The Bank of Japan also seems to have smoothly transitioned to such a system. In both cases, negative rates were signalled in advance to give participants time to adjust.

Although a varying required reserves mechanism has not yet been implemented, we believe that it would entail similar transition costs and that it is feasible too. It has the advantage that it just requires the return on required reserves to be above the return on cash. In contrast, the tiered remuneration with varying thresholds implies some central bank policy rates are constrained by the return on cash, such as the central bank lending rate and the deposit rate on exempted reserves. Therefore, although these systems provide some flexibility for the overnight interbank rate to trade below the return on cash, it is not a panacea in that these other central bank rates are constrained, which could also have implications for central bank profitability.

We analyze tiered remuneration using a static model of monetary policy implementation. Our model results would extend to a dynamic model so long as the exemption threshold changes dynamically and changes to the threshold become permanent. That is, since participants will permanently earn the return on cash after reserves are converted to cash, the threshold needs to be permanently changed after each cash conversion such that the penalty for converting to cash is also permanent.

Finally, our model is focused on the monetary policy implementation aspects of negative interest rates. The model does not consider the liability side of bank balance sheets since this is considered beyond scope. For example, retail depositors could withdraw their deposits if interest rates become too negative. This withdrawal could also influence the banks' demand for cash from the central banking sector as well as the compensation banks pay on their deposit liabilities. This could lower the spread banks earn (Jobst and Lin, 2016). We leave these questions for future work.

[8]Participants in Switzerland that have to fulfill reserve requirements are subject to a dynamic exemption threshold, while those that do not are subject to a fixed threshold (Bech and Malkhozov (2016)).

References

Afonso, G., Lagos, R., 2015. Trade dynamics in the market for federal funds. Econometrica 83, 263–313.

Agarwal, R., Kimball, M., 2015. Breaking through the zero lower bound. IMF Working Paper.

Armenter, R., Lester, B., 2015. Excess reserves and monetary policy normalization. Federal Reserve Bank of Philadelphia Working Paper No. 16-33 .

Banerjee, R., Mio, H., 2014. The impact of liquidity regulation on banks. Bank for International Settlements Working Paper No. 470 .

Bech, M., Keister, T., 2013. Liquidity regulation and the implementation of monetary policy. Bank for International Settlements Working Paper No. 432 .

Bech, M., Klee, E., 2011. The mechanics of a graceful exit: Interest on reserves and segmentation in the federal funds market. Journal of Monetary Economics 58, 415–431.

Bech, M., Malkhozov, A., 2016. How have central banks implemented negative policy rates? BIS Quarterly Review.

Bech, M., Monnet, C., 2013. The impact of unconventional monetary policy on the overnight interbank market, in: Liquidity regulation and the implementation of monetary policy, Liquidity and Funding Markets, RBA Annual Conference Volume. Reserve Bank of Australia.

Bindseil, U., 2001. Central bank forecasts of liquidity factors: quality, publication and the control of the overnight rate. European Central Bank Working Paper No. 70 .

Bonner, C., Eijffinger, S., 2012. The impact of the LCR on the interbank money market. De Nederlandsche Bank Working Paper No. 364 .

Borio, C., 1997. The implementation of monetary policy in industrial countries: A survey. Bank for International Settlements Economic Papers No. 47 .

Canzoneri, M., Cumby, R., Diba, B., 2017. Should the federal reserve pay competitive interest on reserves? Journal of Money, Credit and Banking 49, 663–693.

Goodfriend, M., 2002. Interest on reserves and monetary policy, in: Conference on Financial Innovation and Monetary Transmission, Federal Reserve Bank of New York Economic Policy Review, pp. 77–84.

Goodfriend, M., 2016. The case for unencumbering interest rate policy at the zero bound, in: Federal Reserve Bank of Kansas City's 40th Economic Policy Symposium, Jackson Hole, WY, August.

Holthausen, C., Monnet, C., Würtz, F., 2008. Implementing monetary policy without reserve requirements, in: Paper for the Financial Seminar Series, Financial Department, Goethe University, May 27th.

Ihrig, J., Meade, E., Weinbach, G., 2015. Monetary policy 101: A primer on the fed's changing approach to policy implementation. Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2015-047 .

Jobst, A., Lin, H., 2016. Negative Interest Rate Policy (NIRP): Implications for Monetary Transmission and Bank Profitability in the Euro Area. International Monetary Fund.

Kashyap, A., Stein, J., 2012. The optimal conduct of monetary policy with interest on reserves. American Economic Journal: Macroeconomics 4, 266–282.

Keister, T., Martin, A., McAndrews, J., 2008. Divorcing money from monetary policy. Economic Policy Review 14, 41–56.

Klee, E., Senyuz, Z., Yoldas, E., 2016. Effects of changing monetary and regulatory policy on overnight money markets. Board of Governors of the Federal Reserve System Finance and Economics Discussion Series 2016-084 .

Martin, A., McAndrews, J., Palida, A., Skeie, D., 2013. Federal reserve tools for managing rates and reserves. Federal Reserve Bank of New York Staff Report No. 642 .

Poole, W., 1968. Commercial bank reserve management in a stochastic model: Implications for monetary policy. Journal of Finance 23.

Rezende, M., Styczynski, M.F., Vojtech, C., 2016. The Effects of Liquidity Regulation on Monetary Policy Implementation. Technical Report. Working paper, Federal Reserve Board, May.

Swiss National Bank, 2014. Swiss national bank introduces negative interest rates. Press Release, December 18 .

Viñals, J., Gray, S., Eckhold, K., 2016. The broader view: The positive effects of negative nominal interest rates. iMFdirect blog, April 10 .

Whitesell, W., 2006. Interest rate corridors and reserves. Journal of Monetary Economics 53, 1177–1195.

Whittall, C., Goldfarb, S., 2016. The black hole of negative rates is dragging down yields everywhere. Wall Street Journal, July 10 .

Williamson, S., 2015. Interest on reserves, interbank lending, and monetary policy. Federal Reserve Bank of St. Louis Working Paper 2015-024A .

Witmer, J., Yang, J., 2016. Estimating Canada's effective lower bound. Bank of Canada Review 2016, 3–14.

Appendix A: Modeling tiered reserve rates in a required reserve framework

Modeling tiered reserve rates using the framework developed in section 2 is straightforward. As discussed above, the central bank sets each bank's individual reserve requirement, K^i, to adjust for that bank's cash conversions:

$$K^i = \bar{K} - T^i \tag{.1}$$

Recall that in equation 4, ϵ^i_K is the threshold relative to ϵ^i that determines whether or not central bank borrowing is required. With K^i defined as above:

$$\begin{aligned}\epsilon^i_K &\equiv R^i + \Delta^i - T^i - (\bar{K} - T^i) \\ &= R^i + \Delta^i - \bar{K}\end{aligned}$$

Now, because the reserve requirement is lowered an amount equal to cash conversions, converting to cash doesn't lower the threshold after which central bank borrowing is required. Inserting the new reserve requirement into individual banks' profit function in equation 7 yields a slightly modified expected profit function:

$$\mathrm{E}[\pi^i] = r_B B^i - r_D(D^i - \epsilon^i) + r_C(C^i + T^i) - r_\Delta \Delta^i + r_K(\bar{K} - T^i) + r_R \epsilon^i_K - (r_X - r_R)\int_{\epsilon^i_K}^{\infty} (\epsilon^i - \epsilon^i_K)\mathrm{d}\epsilon^i \tag{.2}$$

Maximizing expected profit subject to the same constraint that $T^i \geq 0$ yields two equilibrium outcomes:

$$\begin{aligned}r_\Delta &= r_R + (r_X - r_R)[1 - G(R - \bar{K})] \\ r_C + \lambda &= r_K\end{aligned}$$

In the first, $\lambda > 0$ and $T = 0$. This equilibrium can exist only if $r_K > r_C$. Recall that r_K, the return paid on required reserves, is controlled by the monetary authority. Therefore, the monetary authority can ensure that the bank-specific reserve requirement in equation .1 yields zero cash conversions by also setting $r_K > r_C$. The second equilibrium is when $r_K = r_C$ and banks are thus indifferent for any $T_i \geq 0$. Clearly, this can easily be avoided by ensuring that $r_K > r_C$.

Appendix B: Monetary Policy Implementation with Payment Shocks to Cash and Reserves

In a world where cash is used to settle more payments and reserves are used less (i.e., the volatility of the reserves payment shock decreases), less aggregate reserves are needed for the interbank rate to equal the deposit rate. In the limit, when this payment shock volatility approaches zero, an epsilon positive amount of excess reserves drives the interbank rate to the floor.

Specifically, we adapt the model by considering two deposit shocks - one that affects cash balances (ϵ^i_C) and one that affects reserve balances (ϵ^i_R). Central bank borrowing can now be represented relative to the shock that affects reserves:

$$X^i_R = \max\{0, \epsilon^i_R - \epsilon^i_K\} \tag{.3}$$

$$X^i_C = \max\{0, \epsilon^i_C - (C^i + T^i)\} \tag{.4}$$

Profits will now be affect by both types of shocks:

$$\pi^i = r_C(\mathbf{EC}^i) - r_\Delta \Delta^i + r_K K^i \tag{.5}$$

$$- r_X \mathbf{X^i} = r - R\mathbf{ER}^i + r_B B^i \tag{.6}$$

$$- r_D(D^i - \epsilon^{\mathbf{i}}_{\mathbf{R}} - \epsilon^{\mathbf{i}}_{\mathbf{C}} \tag{.7}$$

where $ER^i = (R^i + X^i_R + \Delta^i - t^i - \epsilon^i_R) - K^i$ and $EC^i = (C^i + T^i + X^i_C - \epsilon^i_C)$. A larger cash-deposit shock will lower deposit balances and cash balances, while a larger reserve-deposit shock will lower deposit balances and reserve balances (or necessitate central bank borrowing). We assume that participants are able to borrow cash from the central bank if the cash payment shock would make cash balances negative. If that were not the case, they would hold large cash balances when interest rates are negative, such that they are large enough to meet the most extreme cash payment shock. In this case, the participant will not need to borrow from the central bank if hit with a large cash payment shock, and the marginal benefit of holding cash is r_C rather than the benefit calculated below.

With these two kinds of payment shocks, expected profit is:

$$
\begin{aligned}
\mathrm{E}[\pi^i] = {} & r_B B^i - r_D D^i - r_\Delta \Delta^i \\
& + r_C (C^i + T^i) \\
& + r_C \int_{-\infty}^{C^i + T^i} (C^i + T^i - \epsilon_C^i) \mathrm{dF}(\epsilon_C^i) \\
& - r_X \int_{c^i = T^i}^{\infty} (\epsilon_C^i - (C^i + T^i)) \mathrm{dF}(\epsilon_C^i) \\
& + r_R \int_{-\infty}^{R^i + \Delta^i - T^i} (R^i + \Delta^i - T^i - \epsilon_R^i) \mathrm{dG}(\epsilon_R^i) \\
& - r_X \int_{R^i + \Delta^i - T^i}^{\infty} (\epsilon_R^i - (R^i + \Delta^i - T^i)) \mathrm{dG}(\epsilon_R^i).
\end{aligned}
$$

Aggregating across all commercial banks and using the fact that $\Delta = 0$, we can write the first order condition for interbank borrowing as

$$0 = -r_\Delta + r_R G(R - T) + r_X (1 - G(R - T))$$

and the first order condition for cash transfers as

$$0 = r_C F(C + T) + r_X (1 - F(C + T) - r_R G(R - T) - r_X (1 - G(R - T)) + \lambda.$$

Rearranging and combining, the two equations can be expressed as:

$$
\begin{aligned}
r_\Delta &= r_X (1 - G(R - T)) + r_R G(R - T)] \\
r_\Delta &= r_X (1 - F(C + T) + r_C F(C + T) + \lambda.
\end{aligned}
$$

Relative to the set-up in the paper, there is a higher benefit to holding cash balances since a payment shock can also decrease cash balances and necessitate central bank borrowing. The marginal benefit thus includes a benefit in the probability that the bank needs to pay the borrowing rate on central bank funds, r_X. If the volatility of the cash payment shock is zero, the equation reverts to that in the paper (assuming positive cash balances).

On the other hand, in the case when the volatility of the reserves payment shock reduces to zero - i.e., when payment shocks are made with cash instead of reserves, only a small amount of excess reserves are necessary to implement a floor system of monetary policy implementation. In this case, the deposit rate and the target rate would be equal.