Working Paper No. 372

On the Negatives of Negative Interest Rates and the Positives of Exemption Thresholds

Aleksander Berentsen, Hugo van Buggenum and Romina Ruprecht

December 2020

On the Negatives of Negative Interest Rates and the Positives of Exemption Thresholds [*]

Aleksander Berentsen

University of Basel and Federal Reserve Bank of St. Louis

Hugo van Buggenum   Romina Ruprecht

Tilburg University   University of Basel

December 16, 2020

Abstract

Major central banks remunerate reserves at negative interest rates and it is increasingly likely that they will keep rates negative for many more years. To study the long run implications of negative rates, we construct a dynamic general equilibrium model with commercial banks funding investment projects and a central bank issuing reserves. Negative rates distort investment decisions resulting in lower output and welfare. These findings sharply contrast the short-run expansionary effects ascribed to negative rate policies by most of the existing literature. Negative rates also reduce commercial bank profitability. Exempting a fraction of reserves from negative rates can resolve profitability concerns without affecting the central bank's ability to control the money market rate. However, exemption thresholds do no mitigate the investment distortions created by negative rates.

JEL Classification: E40, E42, E43, E50, E58.

Keywords: negative interest rate, money market, monetary policy, interest rates

*"From a micro perspective, low rates undermine financial intermediaries'
health by reducing their profitability, impede the efficient allocation of capital
by enabling even the weakest firms to meet debt-service obligations, and may
also inhibit competition by favoring incumbent firms. There is something
unhealthy about an economy in which corporations can profitably borrow and
invest even if the project in question pays a zero return. These considera-
tions suggest that reducing interest rates may not be merely insufficient, but
actually counterproductive, as a response to secular stagnation."*

Anna Stansbury and Lawrence H. Summers (2019).
Whither Central Banking? Project Syndicate Commentary.

1   Introduction

In recent years, the Bank of Japan, the Danmarks Nationalbank, the European Central
Bank, the Swedish Riksbank and the Swiss National Bank have implemented negative
interest rates (NIR) by decreasing the remuneration of central bank reserves into negative
territory. Most of them grant exemptions from NIR. That is, a fraction of reserves is
remunerated at a rate of zero or even a positive interest rate, while the remaining part
is remunerated at the NIR.

This paper presents a dynamic general equilibrium model of a closed economy to
study the implications of NIR-policies for investment decisions, commercial bank prof-
itability and welfare in steady state. Banks are at the center of our model. Funded by
deposits, they invest, hold central bank reserves and borrow or lend in the money market
at the market clearing interest rate.

The economy is characterized by a distribution of investment projects that differ in
size.[1] In every period, each bank draws one project from this distribution and decides
how much to invest. We study under which monetary conditions banks invest too little
or too much compared to the first-best investment quantity.[2]

A key insight of the model is that NIR distorts investment decisions, lowers welfare
and reduces bank profitability. The type of distortion depends on the transmission of
NIR to deposit rates. With an imperfect transmission, banks with small projects invest
too much compared to the first-best quantities in order to avoid the NIR. In contrast,
banks with large investment projects invest too little because the collateral constraints

---

[1]This feature of the model captures the fact that in reality banks finance projects of various sizes
such as a mortgage for a house or a loan for large plant and that such investment opportunities arrive
at random.

[2]The optimal investment quantity for a particular project is the quantity that a social planner would
choose; i.e., the first-best quantity.

bind and NIR decreases the value of the collateral. Imperfect transmission of NIR to deposit rates has strong support in the empirical NIR-literature and is true across the different NIR-currencies. In particular, the literature agrees that retail deposits (the main funding source for many banks) are largely insulated from NIR.[3]

With perfect transmission of NIR to deposit rates, we do not find the over-investment distortion. The only inefficiency is that large investment projects are underfunded because of binding collateral constraints. This distortion is made worse with NIR.

With perfect and with imperfect transmission of NIR to deposit rates, lowering the interest rate on reserves into negative territory unambiguously lowers welfare. The intuition for this result are the distortions discussed above. Steady state analysis is well suited to reflect the current monetary conditions, where negative rates are implemented over a long time horizon. Switzerland, for example, has introduced negative interest rates of -0.75 percent in mid-January 2015. Initially, central banks were hoping that negative rates would only be needed for a short period of time until the economic conditions would allow to normalize rates back to their positive long-term averages. Unfortunately, this has not yet happened and it is more and more likely that central banks will keep rates in negative territory for many more years.

NIR undermines banks' health by reducing their profitability.[4] We extend the baseline model by introducing exemptions from NIR on part of banks' reserves holdings. Exemptions are motivated by the fact that all of the above-mentioned central banks except for the Swedish Riksbank exempt a fraction of reserves from NIR, remunerating it at zero or even a positive interest rate.[5] We find that exemptions improve bank profitability without affecting the central bank's ability to control the money market rate. However, the investment distortions created by NIR, and the resulting negative welfare effects of NIR, are not mitigated.

Our results have important policy implications: First, they highlight the type of distortions that negative interest rates create in steady state. Second, our model provides very little support for the assumption that NIR stimulates aggregate output. Although, overinvestment occurs if the transmission to deposit rates is imperfect, the damaging effect on large invest projects typically outweighs the former effect. In that sense, our paper is related to Brunnermeier and Koby (2019) and Eggertsson et al. (2019), who

---

[3]Imperfect transmission is documented in Eggertsson et al. (2019), Basten and Mariathasan (2018), Demiralp et al. (2019) and Heider et al. (2019).

[4]While the NIR is the relevant remuneration of banks' asset side (reserves and money market lending), it does not (fully) transmit to the banks' liability side (deposits). This decreases the banks' interest rate margins and thus profits.

[5]Monetary policy with a tiered remuneration of reserves schedule is also studied in Boutros and Witmer (2020). However, their focus is on the implication of exemptions on the demand for physical cash.

also show that NIR can be contractionary under certain conditions. Third, exemptions are an effective remedy against declining bank profits, while at the same time leaving the transmission of NIR to money market interest rates unaffected. Fourth, NIR reduces the aggregate value of a country's currency, indicating that it can be used as a tool to dampen the appreciation of a currency. The Swiss National Bank and the Danmarks Nationalbank explicitly introduced NIR to make their respective currencies less attractive and thus to dampen the appreciation pressure.

Our welfare results are at odds with Agarwal and Kimball (2019) who take the stimulating effect of NIR as given. The premise of their analysis is that NIR is an effective instrument to fight recessions. Based on this fallacy, they discuss how central banks can enable deep negative rates. Our paper clearly suggests that reducing interest rates into negative territory may not be merely insufficient to stimulate the economy, but actually counterproductive as recently suggested by Lawrence H. Summers and Anna Stansbury (see the quote). That being said, NIR might be justified for reasons outside of our model, such as to dampen an excessive appreciation of a currency. In this case, there is a trade-off between the policy goal (dampen the appreciation) and the negative investment and welfare effects of NIR identified in our model.

The remainder of this paper is organized as follows. Section 2 discusses the implementation of NIR across the above-mentioned central banks. Section 3 describes the theoretical model. Section 4 discusses the equilibrium in the baseline economy and the key insights of the model. Section 5 discusses an extension with a haircut on reserves used as collateral. Section 6 reviews the literature on NIR and Section 7 concludes. Finally, all proofs are in Appendix B.

2 Implementation of NIR

In what follows, we describe how NIR have been implemented across different central banks, provide evidence how NIR have transmitted to the economy and discuss the reasons why NIR were implemented by the respective central bank. We focus on the implementation of NIR-policies by the Bank of Japan (BOJ), the Danmarks Nationalbank (DN), the European Central Bank (ECB), the Swedish Riksbank (Riksbank) and the Swiss National Bank (SNB).

Before introducing NIR, the five central banks under consideration operated in an environment with excess reserves. Due to excess reserves, monetary policy is characterized by a so-called floor system (see Berentsen et al. (2014)). They all implemented NIR by decreasing the remuneration of reserves or the interest rate applied in reserve-absorbing operations into negative territory.

Table 1 summarizes the key features of NIR-policies for the five central banks. The second column indicates whether NIR-policies have been implemented with exemptions. Except for the Riksbank, all central banks exempt a fraction of reserves from NIR. The third column describes the reserve remuneration. The ECB, the DN and the SNB have implemented a two-tiered system, where part of the reserves are exempt from NIR and earn an interest rate of 0% (SNB (2015), Jørgensen and Risbjerg (2012)). In Sweden, the Riksbank increased the repo rate to 0% in January 2020. Excess reserves absorbed in fine-tuning operations and deposited at the deposit facility are remunerated at -0.1%.[6] The BOJ has adopted a three-tiered system, where a fraction of reserves earns a positive interest rate, a second fraction earns a zero interest rate and a third fraction is negatively remunerated (Bank of Japan (2016)).[7]

The rationale for exemptions seems to be primarily profitability considerations. Ceteris paribus, NIR decrease banks' interest rate margins if they cannot pass-on NIR to all their liabilities. Heider et al. (2019), Dell'Ariccia et al. (2017), Eisenschmidt and Smets (2019), Eggertsson et al. (2019), Zurbrügg (2016) and Basten and Mariathasan (2018) provide ample evidence that NIR are indeed not passed on to all liabilities. Moreover, the fact that all central banks exclude at least minimum reserve requirements from NIR suggests that central banks are reluctant to charge NIR on required reserves holdings, possibly due to legal considerations.[8] Further, exemptions may have been introduced due to central banks' mandate to ensure the functioning of cashless payments.

The fourth column shows the transmission of NIR to the economy. Consensus in the literature is that NIR have transmitted to money market interest rates and fixed-income markets. This is despite the fact that NIR were introduced with exemptions for all central banks except the Riksbank. Wholesale lending and deposit rates are only partly affected and retail deposits are exempt from NIR, so far. This is true for all five central banks discussed (see Bech and Malkhozov (2016), Dell'Ariccia et al. (2017) and Section 6 that discusses the related literature).

Finally, the fifth column shows the reason why NIR were introduced by the different

---

[6]Even though the Riksbank is effectively no longer operating NIR, we still include it in our discussion because the Riksbank adopted NIR for several years before January 2020.

[7]Banks' exemption threshold calculation is linked to the minimum reserve requirements in cases of the SNB, the ECB and to some extent in case of the BOJ (ECB (2019), SNB (2015), Bank of Japan (2016)). In case of the BOJ, the exemption threshold is based on average reserves holdings, reserve requirements and borrowing in BOJ's loan supporting programs. The latter two are considered in the so-called "Macro-Add on Balance", which is remunerated at 0% and the former represents the so-called "Basic Balance", which is remunerated at +0.1%. Reserve holdings exceeding the basic balances and the macro-add on balances are called the "Policy-Rate Balance" and are remunerated at -0.1% (Bank of Japan (2016)). The DN bases its calculation of the exemption threshold on banks' activity in the payment system. Reserves exceeding the threshold are automatically converted into negatively remunerated certificates of deposits with a one-week tenor (Abildgren et al. (2015)).

[8]There are no minimum reserve requirements at the Riksbank.

As of August 2020	Exemptions	Interest rates	Degree of transmission	Objective of NIR
Bank of Japan	Yes	0.1%, 0%, -0.1%	Partial transmission	Inflation and economic activity
Danmarks Nationalbank	Yes	0%, -0.6%	Partial transmission	Exchange rate considerations
European Central Bank	Yes	0%, -0.5%	Partial transmission	Inflation
Swedish Riksbank	No	0%, -0.1%	Partial transmission	Inflation
Swiss National Bank	Yes	0%, -0.75%	Partial transmission	Exchange rate considerations

Table 1: Implementation of NIR across CB

central banks. In case of the ECB and the Riksbank, the introduction of NIR was part of a larger program including quantitative easing measures, with the goal to increase inflation. Also the BOJ introduced NIR to increase inflation and economic activity together with quantitative easing measures (Bank of Japan (2016)). The Danmarks Nationalbank and the SNB introduced NIR to dampen the appreciation pressure on their respective currencies. In Appendix A, we discuss the implementation of NIR and the environment in which NIR were introduced by the respective central banks in more detail.

3   The Model

Our theoretical model is motivated by the implementation of NIR discussed above. Time is discrete and continues forever: $t = 0, 1, ..., \infty$. There are two types of agents: banks and households. Both agent types are infinitely-lived and each of them has measure 1. The focus of our attention will be on the behavior of banks. Besides households and banks, there is also a central bank that is in charge of monetary policy.

We assume that in each period, two perfectly competitive markets open sequentially

(see Figure 1).[9] First, an *investment-money (IM) market*, where production and investment of a capital good takes place, and where banks can borrow and lend reserves. Second, a *settlement market*, where liabilities are settled and a general good is produced and consumed. All goods are perfectly divisible and nonstorable, which means that they cannot be carried from one market to the next. There are two perfectly divisible financial assets: reserves and deposits. Reserves are issued by the central bank and can only be held by banks. Deposits are liabilities of banks that can only be held by households. Banks are wiling to accept liabilities of other banks at par. Deposits held across periods earn an exogenous nominal rate i_d in the settlement market.

Preferences of households are given by the flow utility function

$$\mathcal{U}(k, x, h) = -c(k) + x - h.$$

Here, x and h are consumption and production, respectively, of the general good by the household. Also, k is production of capital goods and for simplicity, we assume $c(k) = k$.

Preferences of banks are given by the flow utility function

$$\mathcal{V}_\varepsilon(k, x, h) = \varepsilon^{1/\alpha} f(k) + x - h,$$

that we interpret as lump-sum dividends paid to the bank's shareholders. Specifically, banks receive returns $\varepsilon^{1/\alpha} f(k)$ from investing k units of capital. The production function satisfies $f(k) = \frac{k^{1-1/\alpha}}{1-1/\alpha}$, with $\alpha > 1$, and ε is an idiosyncratic investment shock. This shock has a continuous distribution $G(\varepsilon)$ with support $[0, \infty]$, and the shock is i.i.d. across banks and is serially uncorrelated. The main purpose of the investment shock is to generate a distribution of reserve holdings across banks and to study how NIR affect investment decisions for different values of ε. Also, x and h are consumption and production, respectively, of the general good by the bank. Finally, banks and households discount future utility at the common rate $\beta = (1 + r)^{-1} < 1$, where r is the time rate of discount.[10]

First-best allocations. Due to the quasi-linear preference structure, production and consumption of general goods are irrelevant for aggregate welfare. First-best allocations however require that the marginal return from an investment project equals the marginal

---

[9]The theoretical model presented in Section 3 shares elements of Berentsen et al. (2014) and Berentsen et al. (2018). Our theoretical framework allows for a more realistic modeling of the effects of NIR on the economy. Furthermore, effects of NIR were not studied in these papers. Other theoretical papers on money markets include Orr and Mellon (1961), Poole (1968), Furfine (2000), Whitesell (2006), Berentsen and Monnet (2008), Afonso and Lagos (2015), Bech and Monnet (2016), Rocheteau et al. (2018).

[10]Since $\mathcal{V}_\varepsilon(k, x, h)$ represents the bank's dividend payments, utility maximization by a bank implies that the bank maximizes its market value subject to earning an expected return r.
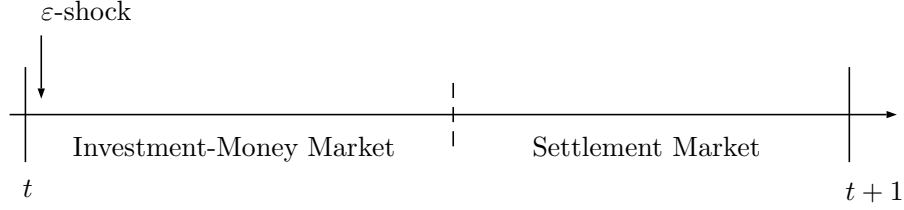
Figure 1: Environment

cost of capital. That is,

$$\varepsilon^{1/\alpha} f'(k) = 1.$$

Solving for k_ε yields

$$k_\varepsilon^* = \varepsilon.$$

Thus, the first-best investment quantities satisfy $k_\varepsilon^* = \varepsilon$ for all ε. Note that $\varepsilon^{1/\alpha} f(k_\varepsilon^*) - k_\varepsilon^* \geq 0$ for all ε. The implication is that from a societal point of view, all projects should be implemented.

In what follows, we consider matters in a market economy. We first discuss the specific structure of the settlement market and the IM market, and the associated decisions of our agents in these two markets. Then, we define equilibrium and characterize welfare.

3.1 Settlement Market

In the settlement market, the general good trades at nominal price P_t^x. We define $\phi_t \equiv 1/P_t^x$ and focus on steady states in which all nominal prices grow at a constant rate between periods.

3.1.1 Central Bank

In the settlement market, the central bank controls the stock of reserves carried into the next period M_{t+1} and chooses an exemption threshold $\overline{M}_t$. The central bank pays nominal interest rate $i_{p,t}$ on reserves below the exemption threshold and $i_{n,t}$ on reserves above the exemption threshold. A bank that carries $\widehat{M}_{\varepsilon,t}$ units of reserves into the settlement market, where we index banks according to the shocks drawn in the IM market, therefore receives a net interest payment

$$\mathcal{P}_{\varepsilon,t} = \begin{cases} i_{p,t}\widehat{M}_{\varepsilon,t} & \text{if } \widehat{M}_{\varepsilon,t} \leq \overline{M}_t \\ i_{p,t}\overline{M}_t + i_{n,t}\left(\widehat{M}_{\varepsilon,t} - \overline{M}_t\right) & \text{if } \widehat{M}_{\varepsilon,t} > \overline{M}_t \end{cases}$$

8

from the central bank. For most parts of our policy analysis, we think of $i_{n,t}$ to be negative (the NIR) and $i_{p,t}$ to be equal to zero to capture the current interest rate policies in NIR countries. However, the equations that follow allow for all cases in which $i_{n,t} \leq i_{p,t}$.

We focus on steady state policies. Let $\gamma = M_{t+1}/M_t$ denote the gross rate at which the stock of reserves grows between periods. The real value of reserves is constant in steady state, so all nominal prices in the economy grow at a gross rate γ between periods. Hence, $\gamma\phi_{t+1} = \phi_t$. We assume that there is a positive initial stock of reserves M_0 and that the exemption threshold grows at the same rate as the stock of reserves.[11] It follows that the remaining policy variables of the central bank are time-invariant interest rates i_n and i_p, and an initial exemption threshold $\overline{M}_0$.

Let $m = \phi_t M_t$ denote the aggregate value of reserves carried into the settlement market, let $\bar{m} = \phi_t \overline{M}_t$ denote the exemption threshold expressed in real terms, and let $\hat{m}_\varepsilon = \phi_t \widehat{M}_{\varepsilon,t}$ denote the real value of reserves carried into the settlement market by a bank that faced investment shock ε. These quantities are all constant in steady state. Given its policies and the distribution of reserves across banks, the central bank's real profits during the settlement market, denoted with π_{CB}, are given by:

$$\pi_{CB} = (\gamma - 1)m - \int_0^\infty [i_p \min\{\hat{m}_\varepsilon, \bar{m}\} + i_n \max\{\hat{m}_\varepsilon - \bar{m}, 0\}]\mathrm{d}G. \tag{1}$$

That means, profits of the central bank equal seignorage revenues minus aggregate interest payments to banks. These profits are distributed to banks and households in a lump-sum fashion. Denote with τ_H (τ_B) a real lump-sum transfer received by all households (resp. banks) during the settlement market, so that

$$\pi_{CB} = \tau_H + \tau_B. \tag{2}$$

3.1.2 Households

To ensure that households' problem is recursive, we characterize value functions for real quantities. Consider a household that enters the settlement market with $\hat{d}$ deposits, expressed in terms of general goods. Let d denote the real amount of deposits that the household carries out of the settlement market, evaluated at the next period's price of general goods. Also, let $W_{IM}(d)$ denote the utility value of entering the next period's IM market with deposits worth d general goods. Defining $\rho_d = 1/(1+i_d)$ and assuming that

---

[11]Since the assets are nominal objects, the government and the central bank can start the economy off with one-time injections of cash M_0

households cannot short-sell deposits, we obtain the following indirect utility function:

$$W_S(\hat{d}) = \max_{x,h,d} \{x - h + \beta W_{IM}(d)\}$$

$$\text{s.t.} \quad x + \gamma d \leq h + \hat{d}/\rho_d + \tau_H, \quad x \geq 0, \quad h \geq 0, \quad \text{and} \quad d \geq 0.$$

Because the household's budget constraint will always hold with equality, we can eliminate x and h to obtain a value function that is linear in $\hat{d}$:

$$W_S(\hat{d}) = \max_{d \geq 0} \{-\gamma d + \beta W_{IM}(d)\} + \hat{d}/\rho_d + \tau_H. \tag{3}$$

Conjecture that $W_{IM}(d)$ is continuously differentiable and concave, and let $W_{IM}^d = \partial W_{IM}/\partial d$. We obtain the following necessary and sufficient condition for optimality of deposits carried out of the settlement market:

$$0 \geq -\gamma + \beta W_{IM}^d(d), \quad \text{with equality if } d > 0. \tag{4}$$

3.1.3 Banks

To ensure that bankers' problem is recursive, we again characterize value functions for real quantities. Consider a bank that enters the settlement market with reserves worth $\hat{m}$ general goods, reserves borrowed from other banks during the preceding IM market worth $\hat{z}$ general goods, and deposits worth $\hat{d}$ general goods. Define $\rho_n = 1/(1 + i_n)$, $\rho_p = 1/(1 + i_p)$, and $\rho_m = 1/(1 + i_m)$, where i_m is the nominal interest rate for borrowing reserves from other banks in the IM market. Let m denote the real amount of reserves that the bank carries out of the settlement market, evaluated at the next period's price of general goods. We impose a non-negativity constraint on m to capture that banks cannot short reserves. Also, let d be the real amount of deposits on the bank's balance sheet at the end of the settlement market, again evaluated at the next period's price of general goods. Finally, let $V_{IM}(m, d|\varepsilon)$ denote the bank's indirect utility function of entering the next IM market with reserves and deposits worth m and, respectively, d general goods, conditional on receiving investment shock ε. For the bank, we obtain the following indirect utility function associated with entering the settlement market:

$$V_S(\hat{m}, \hat{z}, \hat{d}) = \max_{x,h,m} \left\{ x - h + \beta \int_0^\infty V_{IM}(m, d|\varepsilon) dG \right\}$$

$$\text{s.t.} \quad x + \frac{\hat{d}}{\rho_d} + \gamma(m - d) \leq h - \frac{\hat{z}}{\rho_m} + \frac{\min\{\hat{m}, \bar{m}\}}{\rho_p} + \frac{\max\{\hat{m} - \bar{m}, 0\}}{\rho_n} + \tau_B,$$

$$x \geq 0, \quad h \geq 0, \quad \text{and} \quad m \geq 0.$$

10

Because the bank's budget constraint will always hold with equality, we can eliminate x and h to obtain a value function that is linear in $\hat{m}$, $\hat{z}$, and $\hat{d}$:

$$V_S(\hat{m}, \hat{z}, \hat{d}) = \max_{m \geq 0} \left\{ -\gamma(m - d) + \beta \int_0^\infty V_{IM}(m, d|\varepsilon) \mathrm{d}G \right\}$$
$$- \hat{d}/\rho_d - \hat{z}/\rho_m + \min\{\hat{m}, \bar{m}\}/\rho_p + \max\{\hat{m} - \bar{m}, 0\}/\rho_n + \tau_B. \quad (5)$$

Conjecture that $V_{IM}(m, d|\varepsilon)$ is continuously differentiable and concave in m, and let $V_{IM}^m = \partial V_{IM}/\partial m$. We then obtain the following necessary and sufficient condition for optimality of reserves carried out of the settlement market:

$$0 \geq -\gamma + \beta \int_0^\infty V_{IM}^m(m, d|\varepsilon) \mathrm{d}G, \quad \text{with equality if } m > 0. \quad (6)$$

3.2 Investment-Money Market

During the IM market, banks and households can trade capital goods at nominal price P_t^k. Let $p = P_t^k/P_t^x$ denote the relative price of capital goods in terms of general goods, which is constant in steady state. We assume that each household produces the same amount of capital k_s and that each bank has the same customer base. By producing capital, households therefore acquire a portfolio of deposits that is evenly distributed across banks. Besides acquiring capital, banks can lend or borrow reserves in a money market.

In order to introduce a microfoundation for the demand for reserves, we assume that banks are anonymous in the market for capital goods, that banks cannot commit to honor intertemporal promises in this market, and that reserves are the only medium of exchange in this market.

3.2.1 Households

The linearity of $W_S(d)$ implies that $W_S(d) = d/\rho_d + W_S(0)$. Using this property, the household's indirect utility function of entering the IM market with deposits worth d general goods, is given by:

$$W_{IM}(d) = \max_{k_s \geq 0} \left\{ -k_s + [pk_s + d]/\rho_d \right\} + W_S(0). \quad (7)$$

It follows that households are indifferent with respect to the amount of capital that they want to supply if and only if

$$p = \rho_d. \quad (8)$$

Moreover, $W_{IM}(d)$ is continuously differentiable and concave, and $W_{IM}^d(d) = 1/\rho_d$.

3.2.2 Banks

During the IM market, all banks face an inflow pk_s of deposits from households, which the banks take as given. Because these deposits are made by households that produced capital for other banks, the inflow of deposits generates an associated inflow pk_s of reserves. We assume that this inflow takes place just after the IM market has convened.[12]

Consider a bank that has drawn investment shock ε (henceforth the ε-bank). Let m_ε denote the real value of reserves carried out of the IM market by this bank. Note that this quantity is subject to a non-negativity constraint, as reserves cannot be shorted:

$$m_\varepsilon \geq 0. \tag{9}$$

The ε-bank's real reserve holdings at the beginning of the settlement market are then given by $m_\varepsilon + pk_s$. In the money market, the ε-bank can borrow reserves from (or lend reserves to) other banks at a competitive nominal rate i_m. Let z_ε denote the real value of net reserves borrowed in the money market by the ε-bank. We assume that to borrow reserves, banks must pledge collateral. A bank can pledge a fraction $\sigma \in [0,1]$ of reserves carried out the IM market as collateral. Also, a fraction $\theta \in [0,1]$ of the reserves acquired just after the IM has convened can be pledged as collateral. The interpretation of $\sigma < 1$ or $\theta < 1$ is a haircut on the value of reserves pledged as collateral, for example because the bank can abscond with some of these reserves in case it chooses to default on its obligations. Hence,

$$z_\varepsilon \leq \sigma m_\varepsilon + \theta pk_s.$$

Suppose that the ε-bank enters the IM market with m real reserve holdings and d real deposits. When k_ε denotes the amount of capital goods acquired by the ε-bank, we obtain as budget constraint

$$m_\varepsilon + pk_\varepsilon \leq m + z_\varepsilon.$$

Because profit maximizing banks do not leave resources on the table, this budget constraint must hold with equality. Combining with the borrowing constraint, we find

$$pk_\varepsilon + m_\varepsilon(1 - \sigma) - m \leq \theta pk_s. \tag{10}$$

Exploiting the linearity of $V_S(\hat{m}, \hat{z}, \hat{d})$ in Equation (6), the ε-bank faces the following

---

[12]We find that the timing of this inflow is irrelevant for our results.

indirect utility function:

$$
V_{IM}(m, d|\varepsilon) = \max_{k_\varepsilon, m_\varepsilon} \left\{ \begin{array}{l} \varepsilon^{1/\alpha} \dfrac{k_\varepsilon^{1-1/\alpha}}{1 - 1/\alpha} + V_S(0) - \dfrac{d + pk_s}{\rho_d} - \dfrac{m_\varepsilon + pk_\varepsilon - m}{\rho_m} \\[2ex] + \dfrac{\min\{m_\varepsilon + pk_s, \bar{m}\}}{\rho_p} + \dfrac{\max\{m_\varepsilon + pk_s - \bar{m}, 0\}}{\rho_n} \end{array} \right\}, \qquad (11)
$$

subject to (9) and (10). We can ignore a non-negativity constraint for k_ε, as this constraint will never bind.

Let μ_ε and λ_ε be the Lagrange multipliers associated with (9) and (10), respectively. First-order conditions for the ε-bank are then given by

$$k_\varepsilon : \quad 0 = (\varepsilon/k_\varepsilon)^{1/\alpha} - p(1/\rho_m + \lambda_\varepsilon) \qquad (12)$$

$$m_\varepsilon : \quad 0 \geq -1/\rho_m + \mathcal{I}_+/\rho_n + (1 - \mathcal{I}_+)/\rho_p - (1 - \sigma)\lambda_\varepsilon + \mu_\varepsilon \qquad (13)$$

$$0 \leq -1/\rho_m + \mathcal{I}_-/\rho_n + (1 - \mathcal{I}_-)/\rho_p - (1 - \sigma)\lambda_\varepsilon + \mu_\varepsilon \qquad (14)$$

where

$$
\mathcal{I}_+ = \begin{cases} 1 & \text{if } m_\varepsilon + pk_s - \bar{m} \geq 0 \\ 0 & \text{if } m_\varepsilon + pk_s - \bar{m} < 0 \end{cases} \quad \text{and} \quad \mathcal{I}_- = \begin{cases} 1 & \text{if } m_\varepsilon + pk_s - \bar{m} > 0 \\ 0 & \text{if } m_\varepsilon + pk_s - \bar{m} \leq 0 \end{cases}.
$$

In Equations (13) and (14), we take into account that a small increase in m_ε may imply that a bank becomes subject to NIR. Equation (13) imposes that a marginal increase in m_ε should not make the bank better off, with $\mathcal{I}_+ = 0$ if the bank remains exempted from NIR for a marginal increase m_ε and $\mathcal{I}_+ = 1$ otherwise. Analogously, Equation (14) imposes that a marginal decrease in m_ε should not make the bank better off, with $\mathcal{I}_- = 1$ if a marginal decrease in m_ε leaves the bank subject to NIR and $\mathcal{I}_- = 0$ otherwise. It can be verified that $V_{IM}(m, d|\varepsilon)$ is concave in m and continuously differentiable in m, d, and ε. Also, $V_{IM}^m(m, d|\varepsilon) = (\varepsilon/k_\varepsilon)^{1/\alpha}/p$.

3.3 Equilibrium and Welfare

Having derived agents' optimal decisions given prices, we can now define what constitutes an equilibrium in our decentralized economy. We focus on steady state equilibria with a strictly positive demand for reserves. In equilibrium, the aggregate demand for capital goods by banks should equal the production of capital goods by households. Moreover, the aggregated net amount of reserves borrowed by banks in the IM market should equal zero. Together with agents' optimal decisions, we find:

13

Definition 1 *Given policy* $(\gamma, \rho_p, \rho_n, \overline{M}_0)$ *and an initial supply of reserves* M_0, *equilibrium is a tuple of real quantities* $\{m, d, k_s, k_\varepsilon, m_\varepsilon\}$ *and a pair of prices* $\{p, \rho_m\}$ *such that:*

1. *Markets clear:* $k_s = \int_0^\infty k_\varepsilon \mathrm{d}G$ *and* $0 < m = \int_0^\infty m_\varepsilon \mathrm{d}G + pk_s$.

2. *Banks maximize profits: with* $\overline{m}M_0 = m\overline{M}_0$, m *solves* (6), *and* $\{m_\varepsilon, k_\varepsilon\}$ *solves* (11) *subject to* (9) *and* (10).

3. *Households maximize utility:* d *solves* (4) *and* k_s *solves* (7).

Proposition 2 *Steady state welfare* $\mathcal{W}$ *satisfies*

$$(1 - \beta)\mathcal{W} = \int_0^\infty \left[\varepsilon^{1/\alpha} f(k_\varepsilon) - k_\varepsilon \right] \mathrm{d}G.$$

Because households' and bankers' flow utility function is linear in production and consumption of general goods, only capital investment matters for welfare.

4   Equilibrium in a Baseline Economy

In this section, we solve a baseline model characterized by $\sigma = 1$. We see this as a reasonable case, since reserves are record-keeping entries at the central bank and therefore good collateral. That means, there should be no haircut imposed on reserves pledged as collateral.

We start the analysis by noting that households only use deposits as a store of value. Hence, combining the first-order condition for deposits carried out of the settlement market (4) with $W_{IM}^d(d) = 1/\rho_d$, we find that households are willing to carry deposits out of the settlement market only if $\gamma\rho_d \leq \beta$. When $\gamma\rho_d < \beta$, households want to carry infinitely many deposits out of the settlement market, as the real return earned by deposits exceeds households' rate of time preference. We therefore need $\gamma\rho_d \geq \beta$ for equilibrium existence. Without loss, we can then focus on equilibria in which households acquire deposits only when they produce capital during the IM market.

Investment Market.   For an equilibrium in the investment market, we need that demand for capital goods equals supply of capital goods. This requires that Equation (8) holds, as otherwise households do not want to supply any capital goods (i.e., when $\rho_d > p$) or want to supply infinitely many capital goods (i.e., when $\rho_d < p$). Using $p = \rho_d$ in banks' first-order condition for capital investment (12), we obtain the following:

Lemma 3 *There exist a critical value ε' which uniquely solves*

$$\varepsilon' = \left(\frac{\rho_d}{\rho_m}\right)^\alpha \frac{m}{\rho_d} + \theta\left[\int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \int_{\varepsilon'}^\infty \varepsilon' \mathrm{d}G\right], \tag{15}$$

such that the quantities of capital invested satisfy

$$k_\varepsilon = \begin{cases} \varepsilon\,(\rho_m/\rho_d)^\alpha & \text{if } \varepsilon \leq \varepsilon' \\ \varepsilon'\,(\rho_m/\rho_d)^\alpha & \text{if } \varepsilon > \varepsilon' \end{cases}. \tag{16}$$

The intuition for Lemma 3 is as follows. Banks determine how much of their disposable reserves to use for acquiring capital at relative price ρ_d and how much reserves to lend or borrow in the money market at relative price $1/\rho_m$. Ideally, banks equate the marginal product of capital investment $(\varepsilon/k_\varepsilon)^{1/\alpha}$ to the opportunity cost of capital investment ρ_d/ρ_m. Banks with productivity $\varepsilon \leq \varepsilon'$ are able to do so. These banks' unconstrained amount of investment, $k_\varepsilon = \varepsilon(\rho_m/\rho_d)^\alpha$, is sufficiently small so that it can be financed with $m + \theta\rho_d k_s$ reserves. This is not true for banks with $\varepsilon > \varepsilon'$. These banks exhaust their borrowing capacity to finance capital investment, as the marginal product of investment then still exceeds the opportunity cost.

Money Market. We first note that the money market rate cannot fall short of NIR. Otherwise, banks want to borrow an infinitely large amount of reserves in the money market to earn at least $i_n > i_m$ on these reserves.

Next, we note that clearance of the money market requires $\int_0^\infty z_\varepsilon \mathrm{d}G = \int_0^\infty [m_\varepsilon + \rho_d k_\varepsilon - m]\mathrm{d}G = 0$, so that gross borrowing equals gross lending. Because of the non-negativity constraint on reserves carried out of the IM market (9), in an equilibrium we have

$$\int_0^\infty k_\varepsilon \mathrm{d}G \leq \frac{m}{\rho_d}. \tag{17}$$

Using Equations (15) and (16) from Lemma 3, we can rewrite Equation (17) as:

$$\varepsilon' \geq \underline{\varepsilon}, \quad \text{where} \quad \underline{\varepsilon} = (1+\theta)\left[\int_0^{\underline{\varepsilon}} \varepsilon \mathrm{d}G + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon}\mathrm{d}G\right] \quad \text{and} \quad \underline{\varepsilon} > 0 \ \text{ if } \ \theta > 0. \tag{18}$$

Here, $\underline{\varepsilon}$ is strictly increasing in θ, $\lim_{\theta \to 0} \underline{\varepsilon} = 0$, and $\theta = 0$ implies $\underline{\varepsilon} = 0$.

Equation (18) can best be understood by noting that Equation (15) implies a negative equilibrium relationship between the money market rate and ε'. Since a low money market rate implies a low opportunity cost of capital investment, banks' optimal amount

of unconstrained investment is large when the money market rate is low. Given the real value of reserves carried into the IM market by banks, the measure of banks that face a binding borrowing constraint is therefore high. Since all these banks exhaust their borrowing constraint, demand for reserves in the money market is large. At the flip side of the same coin, there is a small measure of banks that face a slack borrowing constraint. Moreover, because optimal unconstrained capital investment is high, the maximum amount of reserves that unconstrained banks are willing to supply in the money market is low.

When the money market rate is sufficiently low, that means such that $\varepsilon' < \underline{\varepsilon}$, demand for reserves exceeds supply of reserves. To attain an equilibrium, the money market rate will then increase so that the measure of banks facing a binding borrowing constraint reduces. With the money market rate sufficiently high, we have $\varepsilon' > \underline{\varepsilon}$. Demand for reserves in the money market then falls short of the maximum amount of reserves that unconstrained banks are willing to supply. As a result, some unconstrained banks are holding excess reserves; $m_\varepsilon > 0$ for some ε. In turn, as we shall uncover next, excess reserves have important implications for how NIR affects the money market rate.

Define $\bar{m}' = \max\{\bar{m} - \rho_d k_s, 0\}$, which is the minimal amount of reserves that a bank needs to carry out of the IM market to become subject to NIR. Using our first-order conditions for m_ε (13 and 14), we find that:

Lemma 4 *Reserves carried out of the IM market by the ε-bank satisfy:*

$$m_\varepsilon \leq \begin{cases} \bar{m}' & \text{if } i_n < i_m \leq i_p \\ 0 & \text{if } i_m > i_p \geq i_n \end{cases} \quad \text{and} \quad m_\varepsilon \geq \begin{cases} \bar{m}' & \text{if } i_n \leq i_m < i_p \\ 0 & \text{if } i_m \geq i_p \geq i_n \end{cases}. \tag{19}$$

Lemma 4 is based on the fact that if the money market rate falls short of the interest rate on exempted reserves, banks want to carry at least $\bar{m}'$ reserves out of the IM market. Otherwise, when $\bar{m}' > 0$ banks can make a profit by borrowing more reserves in the IM market at rate i_m and holding them at accounts with the central bank that earn $i_p > i_m$. Additionally, only when the money market rate equals NIR, banks are willing to carry more than $\bar{m}'$ reserves out of the IM market. Similarly, only when the money market rate falls short of or equals the interest rate on exempted reserves, banks are willing to carry a strictly positive amount of reserves out of the IM market.

Finally, using that clearance of the investment market implies $\int_0^\infty k_\varepsilon \mathrm{d}G = k_s$, we find:

Proposition 5 *With a competitive money market, there is full pass through of the policy*

rate i_n to the money market rate i_m if $\bar{m} < m$ and

$$\frac{m}{\rho_d} \geq \left(\frac{\rho_n}{\rho_d}\right)^\alpha \left[\int_0^{\bar{\varepsilon}} \varepsilon \mathrm{d}G + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon} \mathrm{d}G\right]. \tag{20}$$

In the proof of Proposition 5, we first show that there is a floor on the equilibrium money market rate:

$$\rho_m \leq \begin{cases} \rho_n & \text{if } \bar{m} \leq m \\ \rho_p & \text{if } \bar{m} > m \end{cases}. \tag{21}$$

What determines the floor in Equation (21), is whether banks can avoid the NIR by redistributing reserves in the IM market. This is the case when the exemption threshold exceeds the aggregate supply of reserves, as banks can then distribute reserves in the money market so that no bank is subject to NIR. Should the money market rate then fall short of i_p, all banks want to enter the settlement market with reserves at or beyond the exemption threshold (see Lemma 4) while this cannot be the case in equilibrium.

Second, we show that whether the floor in Equation (21) is attained, depends on the borrowing constraint and the value of reserves carried into the IM market:

$$\frac{m}{\rho_d} \geq \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\bar{\varepsilon}} \varepsilon \mathrm{d}G + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon} \mathrm{d}G\right], \text{ with } = \text{if } \rho_m < \begin{cases} \rho_n & \text{if } \bar{m} < m \\ \rho_p & \text{if } \bar{m} \geq m \end{cases}. \tag{22}$$

Specifically, when the money market rate exceeds the floor no bank carries reserves out of the IM market. In that case, $\varepsilon' = \underline{\varepsilon}$ and the money market rate is determined by Equation (22), which must hold with equality to ensure that supply of reserves equals demand for reserves. However, when the value of reserves carried into the IM market is large and parameter θ is small (meaning that $\underline{\varepsilon}$ is small), Equation (22) can only hold with equality if the money markets rate drops below the floor derived in Equation (21). In that case, the money market can only clear if some banks carry reserves out of the IM market and we thus obtain an environment with excess reserves. In turn, this requires the money market rate to be at the floor in Equation (21).

Proposition 5 has important policy implications: The exemption threshold can be chosen arbitrarily close to the stock of reserves without affecting the transmission of NIR to the money market rate. After all, as long as $\bar{m} < m$, the money market rate is determined uniquely by the equilibrium value of reserves carried into the IM market. Moreover, an increase of $\bar{m}$ obviously reduces the interest rate payments of banks to the central bank. In the limit as $\bar{m} \to m$, these payments become arbitrarily small and,

hence, negatively remunerated reserves do not affect banks' profitability.[13] We discuss the implications of NIR and exemption thresholds for the equilibrium value of reserves and bank profitability further below.

Equilibrium. We focus on symmetric stationary equilibria with a strictly positive demand for reserves, full pass-through of NIR to the money market rate, and at least some banks that are subject to the NIR. Such equilibria meet the requirements in Definition 1 and exist only if $\bar{m} < m$. Combining Lemma 3 with $V_{IM}^m(m, d|\varepsilon) = (\varepsilon/k_\varepsilon)^{1/\alpha}/\rho_d$ and the first-order condition for reserves carried out of the settlement market (6), we find:

Proposition 6 *For $\bar{m} < m$, a symmetric stationary equilibrium with a positive demand for reserves and full pass-through of the NIR to the money market rate is sufficiently described by an $\varepsilon' \geq \underline{\varepsilon}$ that solves:*

$$\frac{\gamma \rho_n}{\beta} = \int_0^{\varepsilon'} \mathrm{d}G + \int_{\varepsilon'}^\infty \left(\varepsilon/\varepsilon'\right)^{1/\alpha} \mathrm{d}G. \tag{23}$$

The LHS of Equation (23) captures banks' ex-ante cost of financing capital investment, taking into account that this requires holding reserves. The RHS of Equation (23) governs banks' expected marginal return on capital investment. To equate the costs of carrying reserves to the expected marginal benefits, ε' is determined endogenously. All equilibrium quantities and prices can then be derived as follows. First, Proposition 5 yields $\rho_m = \rho_n$. Second, the relative price of capital p is given by Equation (8); $p = \rho_d$. Third, Lemma 3 yields the real value of reserves m and capital investment k_ε for each ε-bank. Fourth, supply of capital goods satisfies $k_s = \int_0^\infty k_\varepsilon \mathrm{d}G$. Fifth, reserve holdings at the end of the IM market m_ε are given by Lemma 4.

Because the RHS of Equation (23) is strictly decreasing in ε', an equilibrium with full pass-through of NIR to the money market rate is characterized by a unique ε' and exists if and only if

$$1 \leq \frac{\gamma \rho_n}{\beta} \leq \int_0^{\underline{\varepsilon}} \mathrm{d}G + \int_{\underline{\varepsilon}}^\infty (\varepsilon/\underline{\varepsilon})^{1/\alpha} \mathrm{d}G. \tag{24}$$

Equation (24) imposes a lower bound on i_n, below which the NIR no longer passes through to the money market rate. The reason is that the equilibrium value of reserves carried into the IM market then becomes sufficiently low so that the money market rate cannot be at the floor in Equation (21). The equilibrium money market rate then

---

[13]It is important to note that here, we focus on banks' reserves holdings. In reality, reserves (and money market lending) only represents a sub-set of banks' investment opportunities and, depending on the transmission of NIR, they are also subject to the NIR and thus matter for bank profitability.

satisfies

$$\frac{\gamma \rho_m}{\beta} = \int_0^{\bar{\varepsilon}} \varepsilon \mathrm{d}G + \int_{\underline{\varepsilon}}^{\infty} (\varepsilon/\underline{\varepsilon})^{1/\alpha} \, dG, \tag{25}$$

which together with $\varepsilon' = \underline{\varepsilon}$ pins down allocations. In such a regime, small changes in NIR do not affect allocations.

4.1   Effects of NIR in Baseline Economy

In this section, we discuss the effects of NIR for investment decisions and welfare. A key insight of the model is that the effects of NIR depend on the transmission of the money market rate i_m to deposit rate i_d. We intentionally do not provide a theory about how the deposit rate is determined since empirically we observe a wide range of deposit rate behavior.[14] Rather, we provide results for two competing assumptions. We distinguish between perfect transmission and imperfect transmission. Under perfect transmission, we assume that the deposit rate adjusts immediately to a change in i_m so that at any point of time $i_d = i_m$. Under imperfect transmission, we assume that the deposit rate does not adjust to a change in i_m. Furthermore, we distinguish between two cases of imperfect transmission. In the first case $i_d > i_m$. We refer to this case as the NIR case because we observe empirically in all NIR countries that the deposit rates are above the money market rates. In the second case $i_d < i_m$. We refer to this case as the US case because we observe for the US that the federal funds rate is positive while the deposit rates are at zero.

To guide our discussion, first observe that NIR only affects allocations when there is pass-through of NIR to money market rates. Otherwise, the money market rate is determined independently of NIR by Equation (25). Hence, to study effects of NIR we can restrict attention to an equilibrium described by Proposition 6. Lemma 7 characterizes two additional critical values.

Lemma 7 *There exist critical values $\tilde{\varepsilon}$ and $\hat{\varepsilon}$, which satisfy*

$$\tilde{\varepsilon} = \varepsilon' - \theta \left[\int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \int_{\varepsilon'}^{\infty} \varepsilon' \mathrm{d}G \right]$$

$$- \max \left\{ \varepsilon' \frac{\bar{m}}{m} - \left(1 + \theta \frac{\bar{m}}{m} \right) \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^{\infty} \varepsilon' dG \right], 0 \right\}, \tag{26}$$

$$\hat{\varepsilon} = \varepsilon' (\rho_n/\rho_d)^{\alpha}. \tag{27}$$

If and only if $\tilde{\varepsilon} < \varepsilon'$ there are banks that borrow in the money market, and if $\tilde{\varepsilon} < \varepsilon'$ then

---

[14]The transmission of NIR is widely discussed in the literature (see Section 6).

$z_\varepsilon > 0$ for all $\varepsilon > \tilde{\varepsilon}$. If and only if $\hat{\varepsilon} > \varepsilon'$, there exist ε such that $k_\varepsilon > k_\varepsilon^*$ and if $\hat{\varepsilon} > \varepsilon'$, then $k_\varepsilon > k_\varepsilon^*$ if and only if $\varepsilon < \hat{\varepsilon}$.

The implication of Lemma 7 is twofold. First, when $\bar{m} - \rho_d k > 0$ (meaning that banks which leave the IM market with zero reserves are not subject to NIR) and/or $\theta > 0$, lending and borrowing must take place in the money market. Then, $\tilde{\varepsilon} < \varepsilon'$ and banks with $\varepsilon \leq \tilde{\varepsilon}$ lend out reserves and banks with $\varepsilon > \tilde{\varepsilon}$ borrow reserves until they reach the exemption threshold. Second, when $\hat{\varepsilon} > \varepsilon'$ all banks with $\varepsilon < \hat{\varepsilon}$ invest more than first-best quantities.

Perfect transmission: Assume that NIR perfectly transmits to the deposit rate. In this case, $i_d = i_m = i_n$. Figure 2 shows investment quantities under perfect transmission. The $45°$ line in Figure 2 represents the first-best allocation.
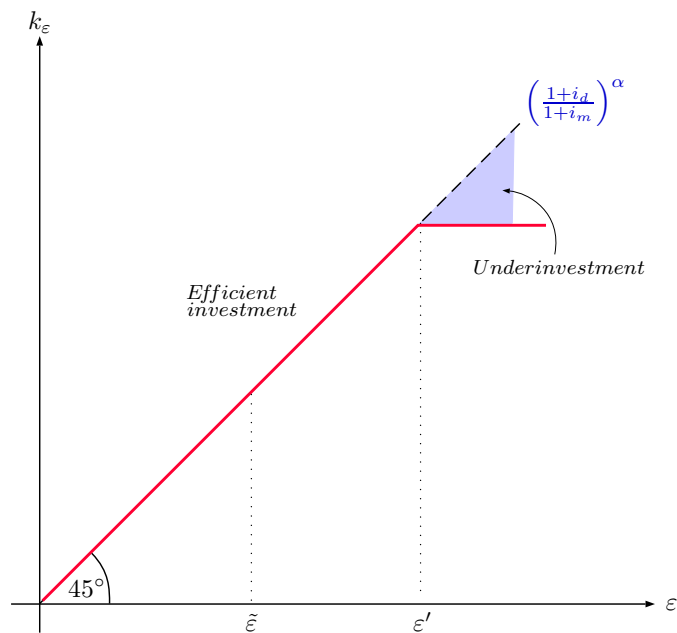
Figure 2: Perfect transmission

The first-best quantities are achieved for all $\varepsilon \leq \varepsilon'$. Banks with $\varepsilon > \varepsilon'$ are constrained and invest less than the first-best quantities. In what follows, we refer to such an investment behavior as underinvestment.

Imperfect transmission (NIR case): Imperfect transmission for the NIR case involves $i_d > i_m = i_n$. As shown in Figure 3, in this case banks with $\varepsilon < \hat{\varepsilon}$ invest more than the first-best quantities. In what follows, we refer to such an investment behavior as

overinvestment. Banks with $0 \leq \varepsilon \leq \tilde{\varepsilon}$ overinvest, because they cannot pass on the NIR to their depositors. Banks with $\tilde{\varepsilon} \leq \varepsilon \leq \hat{\varepsilon}$ overinvest because they can borrow reserves at a lower rate i_m and deposit them at the central bank at interest rate i_p. Among the banks that overinvest, those with $\varepsilon \leq \varepsilon'$ are unconstrained and those with $\varepsilon' < \varepsilon < \hat{\varepsilon}$ are constrained by their reserve holdings. Finally, banks with $\varepsilon > \hat{\varepsilon}$ are constrained and underinvest. A bank with investment shock $\varepsilon = \hat{\varepsilon}$ is constrained, but nevertheless invest the first-best quantity (see Figure 3).
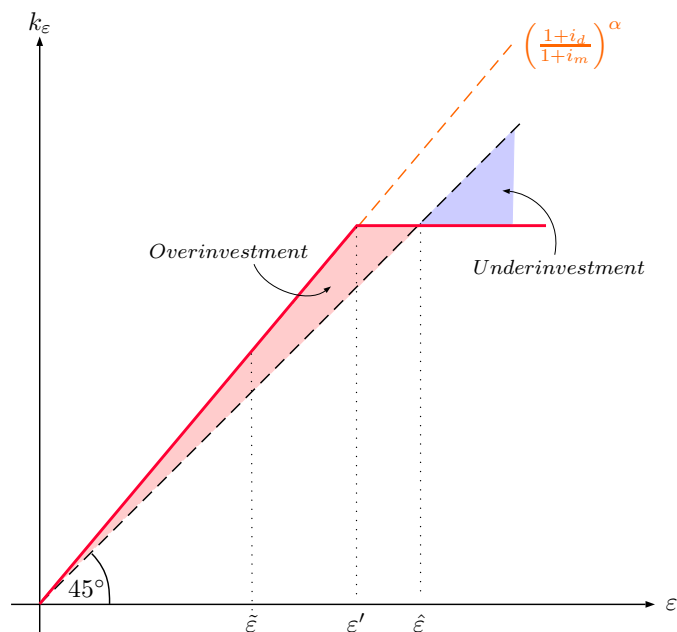

Figure 3: Imperfect transmission: NIR

Imperfect transmission of NIR to bank deposit rates is widely discussed in literature, underlining the relevance of this case.[15] For this case, our model predicts overinvestment, which is confirmed by empirical evidence for NIR-countries. Our interpretation of the empirical manifestation of overinvestment is also emphasized in the quote by Stansbury and Summers (2019) (see the introduction).

Imperfect transmission (US case):   Imperfect transmission for the US case involves $i_d < i_m = i_n$. As shown in Figure 4, in this case all banks underinvest because they receive a higher rate on reserves than the interest rate on deposits. Here, banks with $0 \leq \varepsilon \leq \varepsilon'$ are unconstrained and banks with $\varepsilon > \varepsilon'$ are constrained by their reserve holdings.

---

[15]See for example Dell'Ariccia et al. (2017), Basten and Mariathasan (2018), Heider et al. (2019), Eisenschmidt and Smets (2019), Demiralp et al. (2019) or Eggertsson et al. (2019)
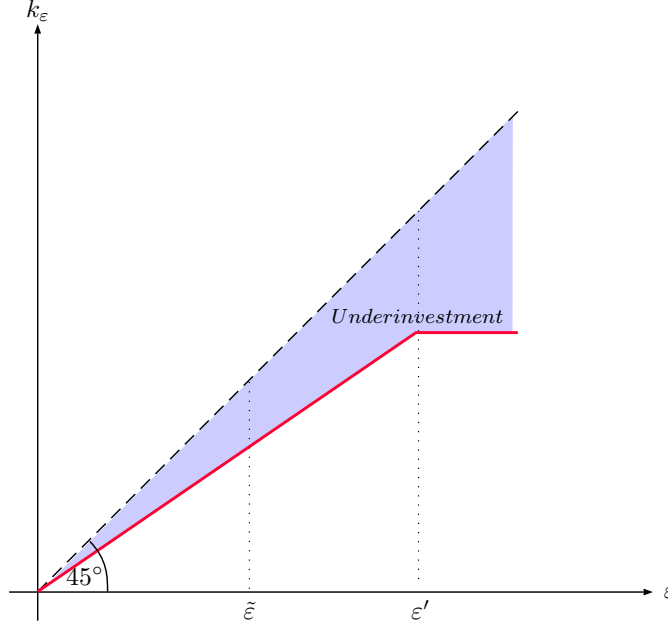
Figure 4: Imperfect transmission: US

4.2 Welfare effects from lowering NIR in Baseline Economy

In what follows, we discuss the welfare effects of a decrease in i_n. Recall that if NIR is fully passed-through to the money market rate, any change in i_n is directly reflected in a change in i_m. In this section, we focus on this case.[16] In order to establish the welfare effects of NIR, we first show how lowering i_n affects all thresholds and investment quantities with pass-through of NIR to money market rates.

Proposition 8 *A decrease in i_n decreases thresholds $\tilde{\varepsilon}$, ε', and $\hat{\varepsilon}$. Furthermore, for all $\varepsilon < \varepsilon'$ with perfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} = 0$ and with imperfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} > 0$. Finally, for all $\varepsilon > \varepsilon'$ with both perfect and imperfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} < 0$.*

A decrease in i_n increases the measure of banks that underinvest since all banks with $\varepsilon < \hat{\varepsilon}$ underinvest and $\hat{\varepsilon}$ is increasing in i_n. Similarly, lowering i_n increases the measure of constrained banks and implies that all constrained banks invest less. Further, a decrease in i_n decreases $\tilde{\varepsilon}$. When $\tilde{\varepsilon} < \varepsilon'$ this leads to more borrowing in the money market, since the measure of banks with $\varepsilon > \tilde{\varepsilon}$ increases.

Proposition 9 *Using Proposition 2, we show that the derivative $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n}$ can be broken*

---

[16]If the NIR does not pass-through to the money market rate, then small changes in NIR have no effect on equilibrium allocations.

into three terms as follows:

$$\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = A + B + C, \ where \tag{28}$$

$$A \equiv \rho_d \int_0^{\varepsilon'} (i_n - i_d) \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} \mathrm{d}G,$$

$$B \equiv \int_{\varepsilon'}^{\hat{\varepsilon}} \left[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1 \right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} \mathrm{d}G,$$

$$C \equiv \int_{\hat{\varepsilon}}^{\infty} \left[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1 \right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} \mathrm{d}G,$$

Furthermore, we can show that the welfare effect of a decrease in the NIR with perfect or imperfect transmission is always negative, $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} < 0$.

The term A captures the welfare changes of those banks that are unconstrained. The term B captures the welfare changes of those banks that are constrained and overinvest. Finally, the term C captures the welfare changes of those banks that are constrained and underinvest. The aggregate effect of a decrease in i_n on welfare is always negative. Below, we provide an intuition for this result depending on transmission of money market rates to deposits.

Perfect transmission: Assume that NIR perfectly transmits to the deposit rate. In this case, $i_d = i_m = i_n$. Figure 5 shows the change in welfare under perfect transmission.

The first-best quantities are achieved for all $\varepsilon \leq \varepsilon'$. After a decrease in i_n, the new equilibrium value is ε''. Consequently, more banks are constrained and invest less than the first-best quantities. This is clearly welfare decreasing. This result is also confirmed by inspecting Equation (28). The term A is zero because with perfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} = 0$. The term B is also zero because with perfect transmission $\varepsilon' = \hat{\varepsilon}$. Finally, the term C is negative because $\varepsilon > \hat{\varepsilon}$ and $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$.

Imperfect transmission (NIR case): Imperfect transmission for the NIR case involves $i_n = i_m < i_d$, with i_d fixed while $i_n = i_m$ as there is pass-through of NIR to the money market rate. Here we distinguish between two cases. First, we consider the effect of a decrease in i_n when $i_d = i_m = i_n$ is the initial condition. Second, we consider the effect of a decrease in i_n with $i_d > i_m = i_n$ as initial condition. We summarize the effects for the NIR case in Table 2.

The first case with the initial condition $i_d = i_m = i_n$, is depicted in Figure 6. Unconstrained banks with $0 \leq \varepsilon \leq \varepsilon'$ initially invest the efficient quantity. The decrease in i_n generates a wedge, causing unconstrained banks to overinvest. This effect is clearly
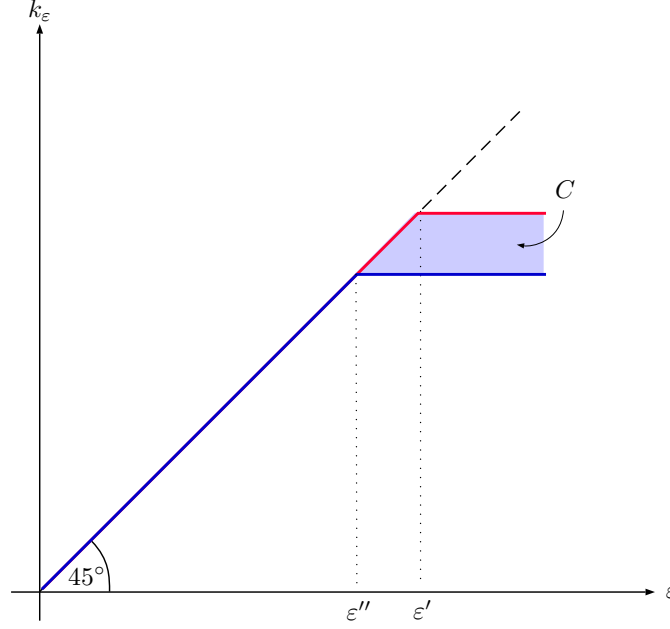
Figure 5: Perfect transmission

negative. Constrained banks with $\varepsilon \geq \varepsilon'$ can invest less after a decrease in i_n because the real value of reserves (expressed in terms of capital goods, i.e. m/ρ_d) decreases. Furthermore, the decrease in i_n also causes a decrease in the critical value ε'. Hence, there are more banks that are constrained and each constrained bank can invest less. This effect is also clearly negative.

Figure 6 suggests that the welfare effect is always negative. This result is confirmed by inspecting Equation (28). The term A is negative because the decrease in i_n implies a negative wedge $i_n - i_d$ and $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} > 0$. The term B is zero because with the initial condition $i_d = i_m = i_n$, $\varepsilon' = \hat{\varepsilon}$. Finally, the term C is negative because $\varepsilon > \hat{\varepsilon}$ and $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$. These effects are summarized in Table 2 in row A' and C'.

The second case with the initial condition $i_d > i_m = i_n$, is shown in Figure 7. Here, banks with $0 \leq \varepsilon \leq \varepsilon'$ initially overinvest. A decrease in i_n lowers the critical values ε' and $\hat{\varepsilon}$ to ε'' and $\hat{\varepsilon}'$, respectively. A decrease in i_n further increases the wedge between i_n and i_d, leading to more overinvestment by unconstrained banks. This effect is clearly negative. Further, the decrease in i_n also decreased the real value of reserves. Hence constrained banks with $\varepsilon > \hat{\varepsilon}$ invest less. This effect is also clearly negative. In the current case, there is also a third effect for banks with $\varepsilon' \leq \varepsilon \leq \hat{\varepsilon}$. As shown in the graph, these banks are constrained and since the decrease in i_n lowers the real value of reserves, these banks can now invest less. Note further, these banks overinvest before
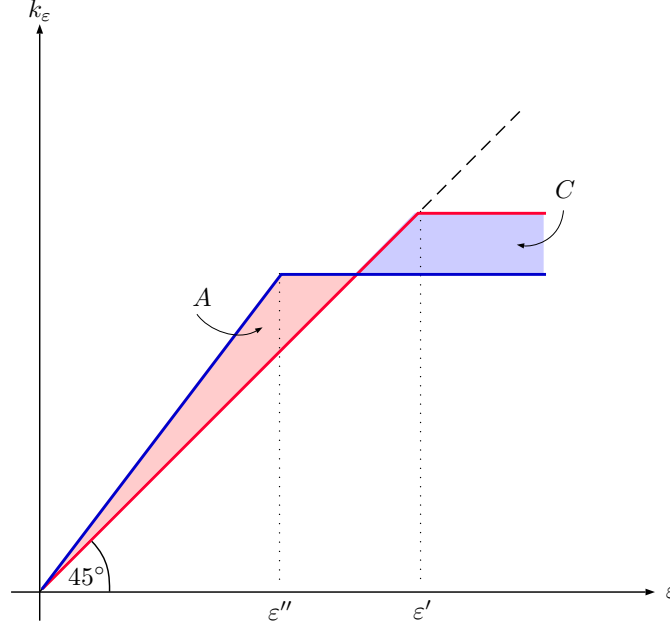
Figure 6: Imperfect transmission: Initial condition $i_d = i_n$

and after the decrease in the NIR. The decrease in the real value of reserves causes these banks to invest a quantity that is lower and therefore closer to the first-best allocation. This has a positive effect on welfare.

The effect on unconstrained banks that overinvest corresponds to term A in Equation (28) and to row A in Table 2. The effect is negative because with imperfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} < 0$ and $i_d > i_m = i_n$. The effect on constrained banks that overinvest corresponds to term B in Equation (28) and to row B in Table 2. This effect is positive because with imperfect transmission $\varepsilon' < \hat{\varepsilon}$. Finally, the effect on unconstrained banks that underinvest corresponds to term C in Equation (28) and to row C in Table 2. This term is negative because of $\varepsilon > \hat{\varepsilon}$ and $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$.

In Proposition 9, we show that a decrease in i_n is always welfare decreasing for imperfect transmission. Intuitively, the positive welfare effect of reduced overinvestment in region B is larger when the interest rate on deposits is high. In the proof of Proposition 9, we confirm our intuition by showing that in an imperfect transmission regime $\mathrm{d}\mathcal{W}(1 - \beta)/\mathrm{d}\rho_\mathrm{n}$ is increasing in the interest rate earned by deposits. However, because deposits can be used as a store of value, there is an upper bound on i_d to have equilibrium existence. This upper bound implies that the negative effects of terms A and C always outweigh the positive effect of term B, leading to an overall negative effect on welfare.
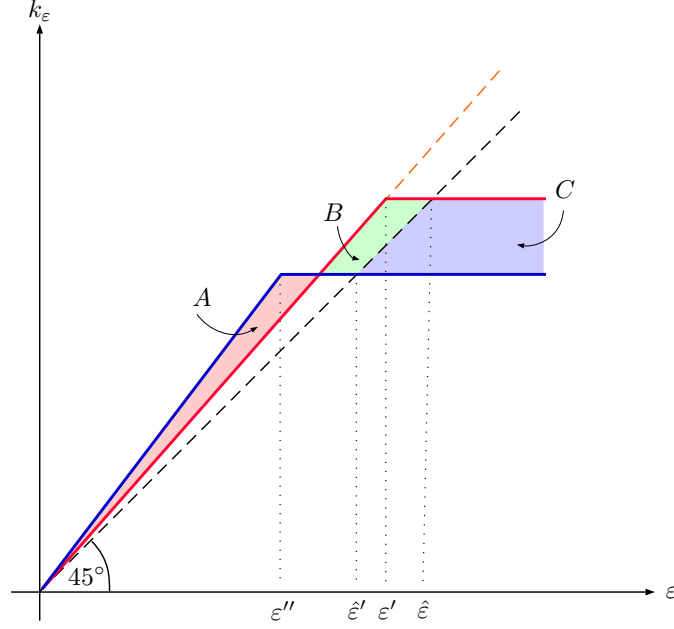
Figure 7: Imperfect transmission: Initial condition $i_d > i_n$

	$i_m < i_d$	overinvest	underinvest	decreasing i_n	$\mathcal{W}$
A	$\varepsilon < \varepsilon'$	$k_\varepsilon > k_\varepsilon^*$		$\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} > 0$	negative
B	$\varepsilon' < \varepsilon \leq \hat\varepsilon$	$k_{\varepsilon'} > k_\varepsilon^*$		$\frac{\mathrm{d}\hat\varepsilon}{\mathrm{d}\rho_\varepsilon} < 0$	positive
C	$\varepsilon > \hat\varepsilon$		$k_{\varepsilon'} < k_\varepsilon^*$	$\frac{\mathrm{d}\hat\varepsilon}{\mathrm{d}\rho_n} < 0$	negative
	$i_n = i_d$				
A'	$\varepsilon < \varepsilon' = \hat\varepsilon$			$\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} > 0$	negative
C'	$\varepsilon > \varepsilon' = \hat\varepsilon$		$k_{\varepsilon'} < k_\varepsilon^*$	$\frac{\mathrm{d}\hat\varepsilon}{\mathrm{d}\rho_n} < 0$	negative

Table 2: Effects of a decrease in i_n in the NIR case

Imperfect transmission (US case): Imperfect transmission for the US case involves $i_n = i_m > i_d$. As shown in Figure 8, all banks underinvest. A decrease in i_n decreases the real value of reserves and therefore decreases ε' to ε''. All banks with $\varepsilon \geq \varepsilon''$ can invest even less with the decrease in the real value of reserves. This effect is clearly negative. However, unconstrained banks with $0 \leq \varepsilon \leq \varepsilon''$ invest more with a decrease in i_n. A decrease in i_n in this scenario, means decreasing the wedge between i_n and i_d and therefore holding reserves becomes relatively less attractive. As a result, unconstrained banks invest more, which has a positive effect on welfare. Proposition 9 shows that a decrease in i_n is always welfare decreasing, implying that the negative effect of term C always outweighs the positive effect of term A. The intuition is again that the positive welfare effect of decreasing i_n depends positively on the interest rate earned by deposits
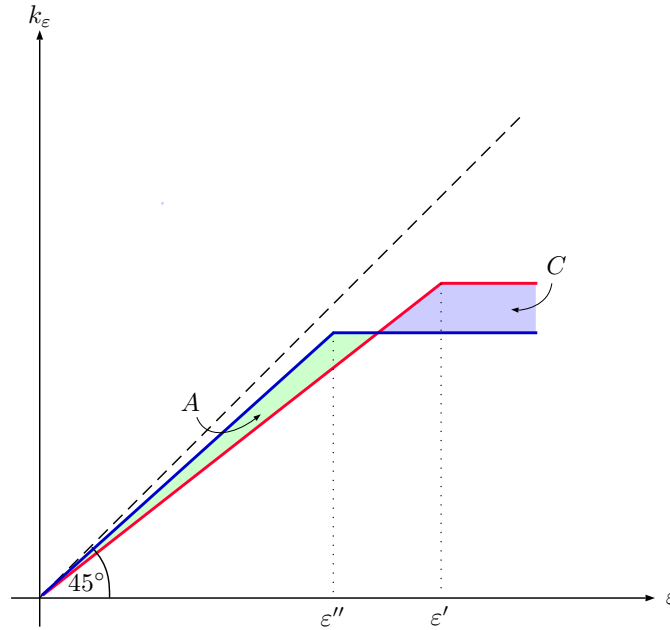
Figure 8: Imperfect transmission: US

i_d, which is bounded from above to have equilibrium existence.

4.3 Effects from lowering NIR on aggregate output and the real value of reserves

In this section, we briefly discuss the effects from lowering i_n on aggregate output and the real value of reserves. Consider first aggregate output, denoted $Q = \int_0^\infty \varepsilon^{1/\alpha} f(k_\varepsilon) \mathrm{d}G$.

Proposition 10 *A decrease in i_n reduces aggregate output.*

Though NIR can lead to more investment, this effect only applies to investment projects with a relatively small efficient scale ($\varepsilon < \varepsilon'$). At the same time, because NIR reduces the amount of investment that constrained banks can undertake, NIR leads to less investment in projects with a relatively large efficient scale ($\varepsilon > \varepsilon'$). Proposition 10 shows that the latter effect dominates the former, so that NIR-policies lead to a reduction in output.

Proposition 11 *A decrease in i_n reduces the real amount of reserves, m.*

This result can be used to justify NIR for exchange rate considerations because the decreased value of reserves (expressed in terms of general goods) can be interpreted as reducing the attractiveness of holding assets in a currency. The SNB and the DN introduced NIR in order to dampen the appreciation of their currencies.

27

4.4 Effects from lowering NIR on bank profitability

In this section, we discuss the effects of a decrease in i_n on bank profitability. Many central bankers are concerned about the effects that NIR have on commercial bank profitability. Bank profits depend on banks' interest payments to the central bank. In Appendix B, we show that real aggregate interest payments that banks receive from the central bank on their reserves, denoted with $\mathcal{P}$, satisfy

$$\mathcal{P} = i_p \bar{m} + i_n (m - \bar{m}).$$

The effect of a decrease in i_n on real aggregate interest payments $\mathcal{P}$ is

$$\frac{\mathrm{d}\mathcal{P}}{\mathrm{d}i_n} = \frac{\mathrm{d}\bar{m}}{\mathrm{d}i_n}(i_p - i_n) + i_n \frac{\mathrm{d}m}{\mathrm{d}i_n} + m - \bar{m}.$$

For $\frac{\mathrm{d}\bar{m}}{\mathrm{d}i_n} > 0$ and $i_n < 0$, a decrease in i_n increases the nominal amount of interest payment to the central bank but also reduces the real value of reserves. The overall effect is ambiguous.

4.5 Optimal Monetary Policy in the Baseline Economy

In what follows, we briefly discuss the optimal monetary policy in the baseline economy. From Equation (23), the Friedman rule implies setting $\rho_n = \beta/\gamma$.

Proposition 12 *The Friedman rule is optimal and implements the first-best allocation for $\rho_n = \rho_d$. For $\rho_d > \rho_n$, the Friedman rule does not implement the first-best allocation.*

This shows that the Friedman rule is the optimal monetary policy and implements the first-best allocation under $\rho_n = \rho_d$. Further, if the central bank implements the Friedman rule, the economy cannot be in the NIR case since equilibrium existence requires $\gamma \rho_d \geq \beta$ and the NIR case would require $\rho_n = \beta/\gamma > \rho_d$. Lastly, for $\rho_n < \rho_d$, the Friedman rule is optimal, but it does not implement the first-best allocation. In this case, all banks underinvest.

5 Equilibrium With a Haircut on Reserves

We now consider equilibrium for $\sigma < 1$ and show that the main insights derived for the baseline economy with $\sigma = 1$ hold true. First, we show that there is a floor on the money market rate given by the NIR.

Lemma 13 *Clearance of the IM market implies $\rho_m \leq \rho_n$.*

We then focus on the relevant case in which there must be at least an ε-bank that ends up with reserves beyond the exemption threshold. To ensure this is indeed the case, following Proposition 5, we assume $\bar{m} < m$. Given this assumption, if the money market rate exceeds the NIR, no bank is willing to carry more than $\bar{m}' = \max\{\bar{m} - pk_s, 0\}$ reserves out of the IM market. Since IM market clearance implies that $pk_s + \int_0^\infty m_\varepsilon dG = m$, we then have $pk_s \geq m - \max\{\bar{m} - pk_s, 0\}$. In turn, this implies $\bar{m} < pk_s$, so all banks leave the IM market without reserves if $i_m > i_n$. Qualitatively, the economy is therefore in one of the following three cases.

No pass-through of NIR. In this case, the money market rate exceeds the NIR so all banks leave the IM market without any reserves. It immediately follows that parameter σ is irrelevant. Hence, banks' optimal investment behavior is given by Lemma 3. Because the money market rate exceeds the floor implied by NIR, the equilibrium money market rate is given by Equation (25). Together with $\varepsilon' = \underline{\varepsilon}$ this pins down allocations. Because we need $\rho_m < \rho_n$, existence of the current case requires

$$\frac{\gamma \rho_n}{\beta} > \int_0^{\underline{\varepsilon}} dG + \int_{\underline{\varepsilon}}^\infty (\varepsilon/\underline{\varepsilon})^{1/\alpha} \, dG.$$

Pass-through of NIR with all banks subject to NIR. In this case, the money market rate equals the NIR and all banks enter the settlement market with reserves greater than the exemption threshold. Evaluating banks' first-order conditions for $\sigma < 1$, the latter requires $\bar{m} - pk_s \leq 0$ so that $m_\varepsilon = 0$ is optimal for all banks. It again follows that parameter σ is irrelevant. Banks' optimal investment behavior is therefore given by Lemma 3 and because the money market rate is now at the floor implied by NIR, ε' is pinned down by Equation (23). Using investment market clearance and Equations (8), (15), and (16), the condition $\bar{m} - pk_s \leq 0$ can be written as:

$$\varepsilon' \leq \left(\frac{m}{\bar{m}} + \theta\right) \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right]. \tag{29}$$

With $\bar{m} \leq m$, Equation (29) is satisfied for $\varepsilon' \leq \bar{\varepsilon}$, where $\bar{\varepsilon}$ depends only on $m/\bar{m}$, θ, and G. Moreover, $\bar{\varepsilon}$ is strictly increasing in $m/\bar{m}$, $\lim_{m/\bar{m} \to 1} \bar{\varepsilon} = \underline{\varepsilon}$, and $\lim_{m/\bar{m} \to \infty} \bar{\varepsilon} = \infty$. Since Equation (18) also needs to be satisfied, existence of the current case requires

$$\int_0^{\bar{\varepsilon}} dG + \int_{\bar{\varepsilon}}^\infty (\varepsilon/\bar{\varepsilon})^{1/\alpha} \, dG \leq \frac{\gamma \rho_n}{\beta} \leq \int_0^{\underline{\varepsilon}} dG + \int_{\underline{\varepsilon}}^\infty (\varepsilon/\underline{\varepsilon})^{1/\alpha}.$$

Pass-through of NIR with some banks subject to NIR. In this case, the money market rate rate equals the NIR, and some banks enter the settlement market with reserves strictly greater than the exemption threshold and some with reserves strictly smaller than the exemption threshold.

From banks' optimization problems and Equation (8), we obtain

$$
k_\varepsilon = \begin{cases}
\varepsilon \left(\frac{\rho_n}{\rho_d} \right)^\alpha & \text{if } \varepsilon \leq \varepsilon' \\
\varepsilon' \left(\frac{\rho_n}{\rho_d} \right)^\alpha & \text{if } \varepsilon' < \varepsilon \leq \varepsilon'', \\
\varepsilon \left(\frac{\rho_n}{\rho_d} \frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha & \text{if } \varepsilon'' < \varepsilon \leq \varepsilon''' \\
\varepsilon''' \left(\frac{\rho_n}{\rho_d} \frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha & \text{if } \varepsilon''' < \varepsilon
\end{cases}
\tag{30}
$$

$$
m_\varepsilon \geq \underline{m}_\varepsilon = \min \left\{ \bar{m} - \rho_d k_s, \frac{m + \theta \rho_d k_s - \rho_d k_\varepsilon}{1 - \sigma} \right\} \quad \text{with } = \text{if } \varepsilon > \varepsilon',
\tag{31}
$$

$$
\varepsilon' = \left[\frac{m - (1-\sigma)\bar{m}}{\rho_d} + (1 - \sigma + \theta) k_s \right] \left(\frac{\rho_d}{\rho_n} \right)^\alpha, \quad \varepsilon'' = \varepsilon' \left(\frac{\rho_n/\rho_p - \sigma}{1 - \sigma} \right)^\alpha,
\tag{32}
$$

$$
\varepsilon''' = \left[\frac{m}{\rho_d} + \theta k_s \right] \left(\frac{\rho_d}{\rho_n} \frac{\rho_n/\rho_p - \sigma}{1 - \sigma} \right)^\alpha.
\tag{33}
$$

Most importantly, we now have three critical values for ε. Like before, all banks with $\varepsilon \leq \varepsilon'$ face a slack borrowing constraint. For all banks with a slack borrowing constraint, the opportunity cost of capital investment is given by $\frac{\rho_d}{\rho_n}$ as additional reserves carried out of the IM market are subject to NIR for these banks. New is that banks with tight borrowing constraints face a trade-off between carrying reserves out of the IM market and investing in productive capital. Specifically, if ε increases beyond ε', banks remain to invest $k_\varepsilon = k_{\varepsilon'}$ but only until the marginal return from capital investment equals the opportunity cost of capital investment when not subject to NIR, which is given by $\frac{\rho_d}{\rho_n} \frac{\rho_n/\rho_p - \sigma}{1 - \sigma}$.[17] That means, until $\varepsilon = \varepsilon''$. Then, as ε increases beyond ε'', it becomes attractive to carry less reserves out of the IM market and to invest more in productive capital. When ε increases beyond ε''', the bank has devoted all available resources to capital investment and remains to invest $k_\varepsilon = k_{\varepsilon'''}$.

Combining Equations (30)-(33) with investment market clearance, money market clearance, and Equation (6), we obtain:

Proposition 14 *Equilibrium with pass-through of NIR and only some banks subject to NIR, is sufficiently described by a tuple $(\varepsilon', \varepsilon'', \varepsilon''')$ that satisfies $0 < \varepsilon' < \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha$,*

---

[17]Banks with $\varepsilon'' < \varepsilon < \varepsilon'''$ carry some reserves out of the IM but are not subject to NIR. In Equations (13) and (14), this implies $\mathcal{I}_+ = \mathcal{I}_- = \mu_\varepsilon = 0$. Therefore, $\lambda_\varepsilon (1 - \sigma) = 1/\rho_p - 1/\rho_m$. Using the latter, $p = \rho_d$, and $\rho_m = \rho_n$ in $p(1/\rho_m + \lambda_\varepsilon)$, i.e. the second part of Equation (12), gives the cost of capital for banks with $\varepsilon'' < \varepsilon < \varepsilon'''$.

$\varepsilon'' = \varepsilon' \left(\frac{\rho_n/\rho_p - \sigma}{1 - \sigma}\right)^\alpha$, and the system of equations

$$\frac{\gamma \rho_n}{\beta} = \int_0^{\varepsilon'} \mathrm{d}G + \int_{\varepsilon'}^{\varepsilon''} \left(\frac{\varepsilon}{\varepsilon'}\right)^{\frac{1}{\alpha}} \mathrm{d}G + \frac{\rho_n/\rho_p - \sigma}{1 - \sigma} \left[\int_{\varepsilon''}^{\varepsilon'''} \mathrm{d}G + \int_{\varepsilon'''}^\infty \left(\frac{\varepsilon}{\varepsilon'''}\right)^{\frac{1}{\alpha}} \mathrm{d}G\right],$$

$$\frac{\varepsilon' - \left[1 - (1 - \sigma)\frac{\bar{m}}{m}\right] \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma}\right)^\alpha}{(1 - \sigma)\left(1 + \theta\frac{\bar{m}}{m}\right)} = \int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \int_{\varepsilon'}^{\varepsilon''} \varepsilon' \mathrm{d}G + \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma}\right)^\alpha \left[\int_{\varepsilon''}^{\varepsilon'''} \varepsilon \mathrm{d}G + \int_{\varepsilon'''}^\infty \varepsilon''' \mathrm{d}G\right].$$

In the proof of Proposition 14, we show that an equilibrium with pass-through of NIR and only some banks subject to NIR is characterized by a unique tuple $(\varepsilon', \varepsilon'', \varepsilon''')$ and exists if and only if

$$1 \leq \frac{\gamma \rho_n}{\beta} \leq \int_0^{\bar{\varepsilon}} \mathrm{d}G + \int_{\bar{\varepsilon}}^\infty (\varepsilon/\bar{\varepsilon})^{1/\alpha} \mathrm{d}G. \tag{34}$$

5.1 Discussion

For low values of θ, and a relatively high value for i_n and the exemption threshold, we find that allocations depend on σ. The reason is that banks which do not carry reserves out of the IM market, then are not subject to NIR in the settlement market. Borrowing reserves in the money market at $i_m = i_n$ to earn $i_p > i_n$ on reserves held with the central bank, then becomes a feasible and profitable investment strategy. With $\sigma < 1$, doing so affects the amount of reserves that a bank can borrow to finance capital investment. As a result, a trade-off between two investment opportunities arises: investment in productive capital and investment in reserves held with the central bank. This trade-off yields different investment behavior than in the baseline economy, as implied by comparing Equation (16) with Equation (30).

Most importantly, even with the trade-off highlighted above present, low ε-banks $(\varepsilon < \varepsilon')$ do not exhaust their borrowing constraints and leave the IM market with reserves that are subject to NIR. These banks therefore face an opportunity cost of capital investment that is governed by the NIR. We therefore find that banks with ε sufficiently low, still overinvest in an imperfect transmission regime with a NIR case $(i_n = i_m < i_d)$. Figure 9 illustrates that when i_n is relatively close to i_d, only banks that are subject to NIR overinvest. When i_n is relatively far away from i_d, Figure 10 illustrates that there are also some banks that are not subject to NIR but that still overinvest.

For the baseline economy with $\sigma = 1$, both with perfect and imperfect transmission of the money market rate to deposits, we know that welfare, output, and the real value of reserves are monotonically decreasing in ρ_n. For the economy with $\sigma < 1$, when ρ_n approaches the Friedman rule, the measure of constrained banks in the IM market
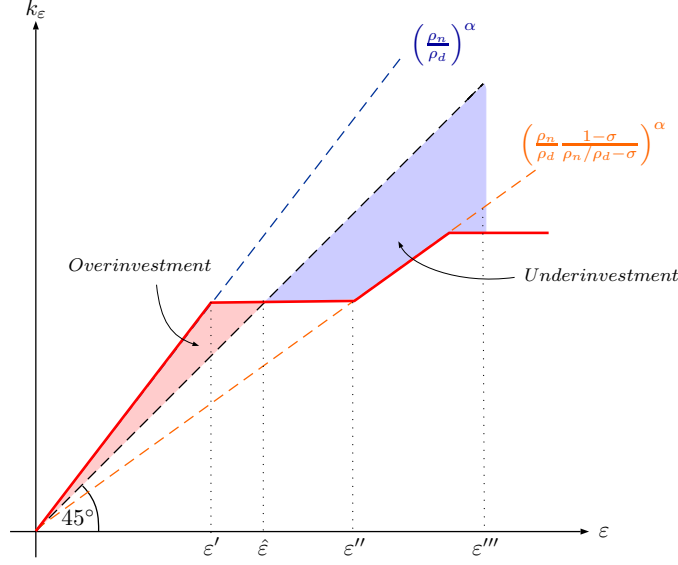
Figure 9: Imperfect transmission: NIR case and $\sigma < 1$ with i_n relatively close to i_d.

approaches zero. Therefore, allocations and welfare become equivalent to that in the baseline model with $\sigma = 1$. Also, when ρ_n exceeds a critical threshold implied by Equation (34), allocations become equivalent to that in the baseline economy. Hence, at least globally, in the economy with $\sigma < 1$ welfare, output, and the real value of reserves are also decreasing in ρ_n. Our most important qualitative results therefore remain unchanged.

An important difference compared to the baseline model is that the exemption threshold can now affect real allocations. Specifically, a higher exemption threshold pushes allocations away from those in the baseline economy because of two reasons. First, it becomes more likely that banks face a trade-off between investment in productive capital and investment in reserves held with the central bank. Second, with a higher exemption threshold that trade-off becomes stronger.

6   Related Literature on Negative Interest Rates

The introduction of NIR by various central banks led to a growing literature that studies NIR as a monetary policy tool. The NIR literature mainly focuses on the transmission of NIR. Consensus among the literature is that NIR have transmitted to money market interest rates and fixed-income markets. Wholesale lending and deposit rates are only partly affected and retail deposits are exempt from NIR, so far. The literature identifies potential risks for bank profitability and financial stability stemming from NIR. These
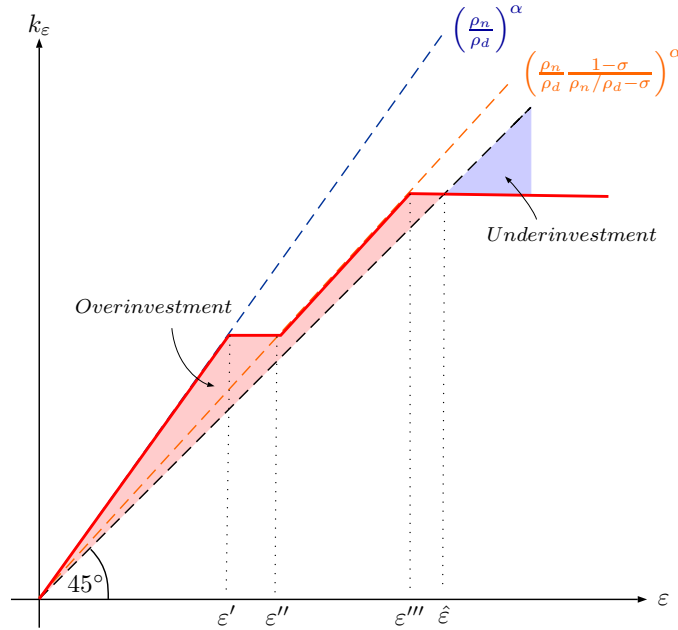
Figure 10: Imperfect transmission: NIR case and $\sigma < 1$ with i_n relatively far from i_d

are particularly relevant when NIR persist over a long period.

Overview articles on NIR include Bech and Malkhozov (2016), Bernhardsen and Lund (2015), Jobst and Lin (2016), Dell'Ariccia et al. (2017), Demiralp et al. (2019) and Jackson (2015). They all discuss the context in which NIR were introduced and how NIR were implemented by respective central banks. Moreover, they discuss the transmission to the money market and beyond and discuss potential side effects associated with NIR.

The theoretical NIR-literature focuses on the optimality of NIR. This paper is most closely related to Dong and Wen (2017), Rognlie (2015), Brunnermeier and Koby (2019), Eggertsson et al. (2019), Ulate (2019) and Porcellacchia (2019). Dong and Wen (2017) analyze optimal monetary policy under NIR in a theoretical framework and find that NIR can be welfare improving because they relax borrowing constraints. Similarly, Rognlie (2015) finds that NIR can be optimal when average output is below potential output. Brunnermeier and Koby (2019) introduce the concept of a reversal interest rate. While an interest rate cut decreases the remuneration on safe assets, it simultaneously increases the valuation of banks' asset holdings. If the first effect dominates the second, there exists a so-called reversal rate, a rate at which monetary policy becomes contractionary. Similarly, Eggertsson et al. (2019) show that NIR need not be expansionary as NIR decrease interest rate margins, which in turn decreases bank profitability. If this effect translates to higher intermediation costs, NIR are contractionary. In contrast, Ulate (2019) studies the effects of NIR in a New Keynesian framework and finds that NIR are

an expansionary monetary policy tool. Also, Porcellacchia (2019) finds that NIR lead to increasing aggregate demand as the incentive to save decreases by studying NIR in a Diamond and Dybvig (1983) framework.

Boutros and Witmer (2020) discuss NIR as a means to overcome the ZLB and specifically address exemption thresholds in their theoretical framework. They show that once the exemption threshold depends on the amount of cash withdrawals, NIR in combination with such an exemption threshold can discourage cash withdrawals and therefore effectively lower the ELB.

The empirical literature focuses on the effects of NIR on bank profitability, bank lending and lending rates. Further topics include the transmission to longer term interest rates, the ELB and investment decisions by firms. In contrast to our paper, Altavilla et al. (2019), who use data on firms in the EU, suggest that NIR are an expansionary monetary policy tool as they find that firms that are subject to negative corporate deposit rates, increase investment.

The effect of NIR on bank profitability is a key topic in the NIR literature. Given that banks earn NIR on (part) of their assets, while only passing it on to part of their liabilities, the banks' interest rate margins deteriorate.[18] This in turn is expected to contribute negatively to the profitability. Turk (2016) however shows that bank profitability in Sweden and Denmark was not negatively affected by NIR. The reason is that banking service fees were increased and wholesale funding costs could be decreased to compensate for the lower interest rate margins from other banking activities. Similar results were found by Basten and Mariathasan (2018) for Swiss banks. However, the authors find that deteriorating interest rate margins were compensated by increasing their exposure towards the interbank market and riskier asset classes. Arseneau (2017) analyzes the effect of NIR on banks in the US using stress test data. He finds that while all banks would anticipate lower profits in a NIR environment, only some would expect lower profits due to net interest rate margins while others would expect higher profits through this channel. For the Eurozone, Demiralp et al. (2019) show that banks holding excess reserves try to circumvent NIR by increasing bank lending and by decreasing wholesale funding. While this evidence points to the fact that banks' have found ways to preserve profits so far, it is an open question how it evolves with a prolonged period of NIR (see for example Jobst and Lin 2016, Dell'Ariccia et al. 2017).

Consensus in the literature is that NIR fully transmitted to money market rates as documented by Bech and Malkhozov (2016), Bernhardsen and Lund (2015), Dell'Ariccia

---

[18]Evidence that that retail deposit rates exhibit a zero lower bound as many banks are reluctant to charge NIR to their retail customers is documented in Eggertsson et al. (2019), Eisenschmidt and Smets (2019), Basten and Mariathasan (2018) and Demiralp et al. (2019).

et al. (2017), Turk (2016), Jackson (2015), Jensen and Spange (2015) and Bräuning and Wu (2017). Also, the literature largely agrees that NIR transmitted to fixed-income markets. Grisse and Schumacher (2017) investigate the effects of short-term interest rate changes on long-term interest rates. Although theory would predict a weakening effect as interest rates approach the effective lower bound, the authors find a stronger effect during the NIR-period than during a zero lower bound period. Similarly, Bräuning and Wu (2017) find that expansionary monetary policy measures had a stronger effect on long-term interest rates during a NIR period than with positive rates. Furthermore, Altavilla et al. (2019) show that NIR transmitted to corporate deposit rates as NIR continued and were lowered further into the negative territory.

The transmission of NIR on bank lending rates is somewhat less clear. Eisenschmidt and Smets (2019) find no evidence that German banks with a higher dependence on retail deposits price their loans differently than banks with a lower dependence on retail deposits. Bräuning and Wu (2017) however find that short-term lending rates decrease in both pre-NIR periods and NIR periods, as a consequence of a decrease in the central bank policy rate. During the NIR period, the authors even find a more distinct reaction on long-term lending rates. Schelling and Towbin (2020) find that banks, which rely more on deposits as a source of funding, offer more generous lending terms in Switzerland. In contrast, Eggertsson et al. (2019) find a decrease of average lending rates in Sweden, but lending rates became more dispersed with the introduction of NIR. Similarly, Amzallag et al. (2019) find that banks, who are more exposed to NIR increased rates on fixed-rate mortgages in Italy.

There is also mixed evidence on the effect of NIR on bank lending volumes in the literature. Bräuning and Wu (2017) find an increase in loan volumes due to NIR. Boeckx et al. (2020) analyze the effect of the ECB's credit-easing policies, which also include NIR - and find increased bank lending volumes due to these policies. However, they do not disentangle the effect arising from NIR and other measures. Ulate (2019) discusses two effects of NIR on bank lending. First, a lower policy rate reduces bank lending rates and thus increases bank lending. Second, a decrease in the net interest rate margin reduces bank profitability and thus lowers the ability of banks to issue loans. Using data on banks in both countries where central banks adopted NIR and where NIR have not been adopted, he finds that the first effect dominates the second effect, suggesting that NIR lead to more bank lending. Also, Bottero et al. (2020) find that NIR lead to increases in bank lending in Italy. These findings stand in contrast to Heider et al. (2019) and Eggertsson et al. (2019), who find that lowering interest rates into negative territory did not increase bank lending in the Euro area and in Sweden. Finally, Heider et al. (2019) find that bank lending shifted towards riskier borrowers in the Euro area.

This is especially true for banks that rely more on deposit funding.

There are a few papers that study the effective lower bound (ELB) with respect to NIR. Lemke and Vladu (2017) use a shadow-rate term structure model to estimate the lower bound, showing that the lower bound became negative after the ECB's second interest rate cut in the NIR territory, but they find the ELB to be above the ECB's negative policy rate at that time. Grisse et al. (2017) analyze how changes in the believed lower bound has affected long-term interest rates. They find that a decrease in the believed lower bound decreases long-term interest rates without the need for a policy rate cut. Related to the ELB discussion, flight to cash issues or the effects on financial stability were discussed by McAndrews (2015). Jensen and Spange (2015) study demand for cash under NIR in Denmark and find no strong increase since the introduction of NIR.

7 Conclusion

In recent years, major central banks have introduced negative interest rates with the aim to increase inflation, stimulate the economy or to dampen the appreciation pressure on the local currency. Most central banks were hoping that negative rates would only be needed for a short period of time until the economic conditions would allow to normalize rates back to their positive long-term averages. Unfortunatley, this has not yet happened and it is more and more likely that central banks will keep rates in negative territory for many more years.

Our paper studies the effects of NIR in steady state in a closed-economy dynamic general equilibrium model. When banks cannot pass on the NIR to their depositors, we identify two distortions. First, overinvestment occurs for small investment projects where banks invest more than the first-best quantity in an attempt to avoid the NIR. Second, investment quantities of large investment projects are too small because collateral constraints bind and NIR decreases the value of collateral (reserves). If the transmission is perfect, we only find that the investment quantities for large projects are too small and that reducing the interest rate on reserves further aggravates this inefficiency. There is a consensus in the literature that NIR is not or only rarely passed on to retail depositors which suggest that the imperfect transmission case is the more realistic case. In both cases, NIR unambiguously decreases welfare. The reason is simply because of the distortions just described.

We also study exemptions from NIR, which are motivated by the fact that most central banks exclude part of the reserves holdings from the NIR, remunerating it at zero or a positive interest rate. We show that exemptions can mitigate the bank profitability

concerns. At the same time, the negative welfare results continue to hold. We also show that NIR lowers the real value of the local currency and therefore is a tool to dampen the appreciation pressure of a currency. The Swiss National Bank and the Danmarks Nationalbank introduced NIR against this background.

NIR-policies are currently a hot topic because many central banks are either already in negative territory or barely in the positive interest rate territory. When these economies enter the next recession, they might decrease policy rates into negative territory — and stay there for a long time. Our paper serves as a warning against adopting uncritically the premise that NIR is an effective tool to fight recession as it is done, for example, in Agarwal and Kimball (2019). Our paper clearly suggests that reducing interest rates into negative territory may not be merely insufficient to stimulate the economy, but actually counterproductive.

Appendix A

Bank of Japan

In early 2016, the BOJ introduced the so-called "Quantitative and Qualitative Monetary Easing with a Negative Interest Rate" program to increase inflation to it's target of 2%. This program entailed the decrease of the interest rate applied to excess reserves to -0.1% and announced a set of loan support programs.[19] At the same time, the BOJ also decided to increase the monetary base and implement an asset purchase program of Japanese government bonds and other financial assets (Bank of Japan 2016).

The BOJ has implemented a three-tiered program for the remuneration of reserves. Average reserve holdings are allocated to the first tier and are referred to as the "Basic Balance", remunerated at a positive interest rate of 0.1%. The second tier called "Macro-Add-on-Balance" is remunerated at 0% and consists of minimum reserve requirements and any amount outstanding from the so-called "Loan Support Program" and the "Funds-Supplying Operation to Support Financial Institutions in Disaster Areas affected by the Great East Japan Earthquake". The third tier is the "Policy-Rate Balance", which is remunerated at a negative rate of -0.1% and contains all reserves not covered in the first two tiers (Bank of Japan 2016). The first and second tier each have an upper bound.[20][21] The implementation of NIR with a three-tiered system is likely due to profit or legal considerations.

By-and-large, NIR transmitted to Japanese Yen overnight money market interest rates. Turnover in the money market has changed significantly in reaction to the introduction with NIR.

---

[19]The loan support programs included the "Loan Support Program", a "Funds-Supplying Operations to Support Financial Institutions in Disaster Areas affected by the Great East Japan Earthquake" and a "Funds-Supplying Operations against Pooled Collateral at zero interest rate" program.

[20]For the first tier the upper bound consist of the difference of the average current account balance or benchmark balance and the required reserves. If the amount computed in the first tier exceeds the upper bound for a financial institution, only the amount covered in the upper bound is remunerated at the positive interest rates. The rest will then count towards the second tier. The upper bound of the second tier is the sum of amount that count towards the loan programs and funds supplying programs as well as some amount related to the benchmark balances. Again, if the amount computed in the second tier exceeds the upper bound, only the amount of the upper bound is remunerated at zero interest rate. The rest will count towards the third tier and thus will be subject to NIR (from `https://www.boj.or.jp/en/statistics/outline/notice_2016/not160616a.pdf`)

[21]The BOJ furthermore also monitors cash holdings of financial institutions to avoid a flight into cash. If the BOJ observes banks with an increase in cash holdings, this amount will be deducted from the second tier and if necessary from the first tier (Bank of Japan 2016).

European Central Bank

The ECB introduced NIR by setting the interest rate on the deposit facility (the lower bound of the channel system) to -0.1% in June 2014. Before that, the deposit facility has been at 0% since 2012. NIR were accompanied by an announcement of targeted longer-term refinancing operations (TLTRO) and a preparation of an expanded purchase program of asset-backed securities. All of these measures were taken to increase inflation and to support bank lending.

This first interest rate cut into the negative territory was followed by four additional interest rate cuts, the last one to date being in September 2019 with a decrease of the deposit facility interest rate to -0.5%.

NIR apply to reserves deposited at the ECB deposit facility as well as all average reserve holdings that exceed the six-fold of minimum reserve requirement held at the current account. Deposits at the deposit facility of the ECB are not subject to exemptions. NIR also applies to government deposits held at the Eurosystem that exceed a certain threshold. Finally, NIR also applies to Eurosystem reserve management services that are not subject to any other interest rate, account balances in TARGET2 (the payment and settlement system of the Euro area), non-Eurosystem overnight deposits held in TARGET2 and other accounts held with Eurosystem central banks that are not subject to other interest rates (ECB 2014).

In September 2019, the ECB implemented a two-tiered system for the adoption of NIR. The six-fold of reserve holdings to fulfill reserve requirements are subject to a zero interest rate, and the rest (excess reserves) is subject to the NIR. The reasons for implementing a two-tiered system is to support the transmission of monetary policy trough the banking system (ECB 2019). Before that, only minimum reserve requirements were exempt from NIR.

In the Euro area, NIR has transmitted to the money market rates. Turnover in the money market seems not to have markedly changed in reaction to the introduction of NIR (Bech and Malkhozov 2016).

Danmarks Nationalbank

The DN has introduced NIR twice in the last few years by lowering the one-week deposit certificate rate into the negative territory. Similar to the SNB, NIR were primarily introduced to alleviate the appreciation pressure on the exchange rate.[22] The first reduction of the one-week certificates of deposit rate was in 2012 and was accompanied by interventions in the foreign exchange market to stabilize the exchange rate of the

---

[22]The DN aims at keeping a nearly fixed exchange rate between the Danish kroner and the Euro, following closely the European Exchange Rate Mechanism II (ERM II).

Danish kroner against the Euro (Jørgensen and Risbjerg 2012). The second reduction of interest rates into the negative territory was in September 2014. The interest rate of the one-week certificates of deposits is currently at -0.6%.

The DN also implemented NIR with a two-tiered system. Reserves held at the central bank are remunerated at the current account rate, which is currently zero. There is a limit on the amount in the current account, determined by the transaction volume in the payment system. Reserve holdings that exceed this limit are subject to the one-week deposit rate (Bech and Malkhozov 2016, Dell'Ariccia et al. 2017, Jørgensen and Risbjerg 2012). Since the introduction of NIR, the DN adjusted the NIR as well as the distribution of current account limits, while keeping the latter constant in the aggregate (Danmarks Nationalbank 2016).

By-and-large, NIR have transmitted to money market rates in Denmark as well. Bech and Malkhozov (2016) report some decreases in the unsecured money market turnover in Denmark.

Sveriges Riksbank

Already in 2009, the Swedish Riksbank introduced NIR, by decreasing the interest rate on the deposit facility to -0.25%. Yet, as the Riksbank conducts daily reserve-absorbing repo fine-tuning operations, the amount of reserves subject to NIR was very small and the NIR did not transmit to money market rates (Jørgensen and Risbjerg 2012, Bech and Malkhozov 2016).

In 2014, the Riksbank entered again the negative territory and introduced NIR by decreasing the repo rate and the deposit facility rate to the negative territory. The introduction of NIR was accompanied by a quantitative easing program. These measures were introduced to increase inflation and to safeguard the role of inflation as nominal anchor (Bech and Malkhozov 2016, Dell'Ariccia et al. 2017, Sveriges Riksbank 2015). In October, the Riksbank increased the deposit rate, followed by a further increase of the repo rate to 0% and the deposit rate to -0.1% in January 2020.

Given the Riksbank's monetary policy implementation framework the relevant NIR is the repo rate. The Riksbank issues one-week debt certificates, which are remunerated at the repo rate. Moreover, to drain reserves prior to the close of business, fine-tuning operations are conducted. The fine-tuning operations are remunerated at 0.10% below the repo rate. Any residual reserve holdings can be deposited at the deposit facility of the central bank which is currently 10 basis points below the repo rate. Thus, the Riksbank effectively no longer operates NIR.

In contrast to all other central banks, the Riksbank effectively operated NIR without

exemptions.[23]

During the NIR period in Sweden, NIR have also transmitted to the money market.

Swiss National Bank

The SNB announced the introduction of NIR of -0.25% on December 18$^{\text{th}}$ 2014 taking effect as of January 21$^{\text{st}}$ 2015 and further decreased the NIR to -0.75% on January 15$^{\text{th}}$ when the SNB discontinued the minimum exchange rate of the Swiss Franc against the Euro. Note, that the announcement on January 15$^{\text{th}}$ also took effect as of January 21$^{\text{th}}$ 2015. NIR were introduced to make investments in Swiss assets relatively less attractive, alleviating the appreciation of the Swiss Franc against the Euro.

The interest rate on excess reserves has been set to -0.75% and has not changed since the introduction in January 2015.

The SNB implemented NIR with a two-tiered system. Reserve holdings above the exemption threshold are remunerated at -0.75% and reserve holdings below the exemption threshold at 0%. Exemptions are calculated according to two methods and are at least CHF 10 millions. The first method is based on a basis component and a cash component. The basis component consists of the moving average of minimum reserve requirements over 36 reference periods multiplied by a threshold factor, which is currently set at 25. The cash component consists of cash holdings during the last reference period and is subtracted from the basis component. This implies that banks cannot hold more cash in order to avoid the NIR. Furthermore, the exemption threshold has to be at least as high as minimum reserve requirements in the last reference period. The second method applies to all financial institutions that are not subject to the first method. These banks have a fixed exemption threshold (SNB 2019).

Also in Switzerland, NIR have transmitted to money market rates. The introduction of NIR with exemptions also caused a sharp increase of money market turnover in Switzerland as banks reallocate reserves from FIs facing NIR to those, which hold reserves below their exemptions.

Other Central Banks

Aside from the five central banks discussed above, Norges Bank, Hungary and the Bulgarian National Bank have also adopted some sort of NIR for some time. These central banks are not in the focus of our analysis because the NIR did not transmit to money markets and beyond.

---

[23]Swedish banks are not subject to minimum reserve requirements. This could partly explain why there was also no need for exemptions in the Swedish case.

Appendix B

Proof of Proposition 2.   In steady state equilibrium, all banks and households enter the IM with reserves worth m general goods and deposits worth d general goods, respectively. By construction, $\mathcal{W} = \int_0^\infty V_{IM}(m, d|\varepsilon)dG + W_{IM}(d)$. By Definition 1, $\{m_\varepsilon, k_\varepsilon\}$ solves (11) subject to (9) and (10), and k_s solves (7). Hence, we can substitute out $V_{IM}(m, d|\varepsilon)$ and $W_{IM}(d)$ to obtain

$$\mathcal{W} = \int_0^\infty \left[\begin{array}{l} \varepsilon^{1/\alpha} f(k_\varepsilon) + (m - m_\varepsilon - pk_\varepsilon)/\rho_m \\ + \min\{m_\varepsilon + pk_s, \bar{m}\}/\rho_p + \max\{m_\varepsilon + pk_s - \bar{m}, 0\}/\rho_n \end{array} \right] dG$$
$$- k_s + V_S(0) + W_S(0).$$

Observe that by construction, reserves carried into the settlement market by a ε-bank satisfy $\hat{m}_\varepsilon = m_\varepsilon + pk_s$. Using IM market clearance conditions $\int_0^\infty k_\varepsilon dG = k_s$ and $m = \int_0^\infty m_\varepsilon dG + pk_s$, we may rewrite the above as

$$\mathcal{W} = \int_0^\infty \left[\varepsilon^{1/\alpha} f(k_\varepsilon) - k_\varepsilon + i_p \min\{\hat{m}_\varepsilon, \bar{m}\} + i_n \max\{\hat{m}_\varepsilon - \bar{m}, 0\} \right] dG$$
$$+ m + V_S(0) + W_S(0).$$

From Definition 1, it follows that m solves (5) and d solves (3). Hence, using (5) and (3), we can substitute out $V_S(0)$ and $W_S(0)$ to obtain:

$$\mathcal{W} = \int_0^\infty \left[\varepsilon^{1/\alpha} f(k_\varepsilon) - k_\varepsilon + i_p \min\{\hat{m}_\varepsilon, \bar{m}\} + i_n \max\{\hat{m}_\varepsilon - \bar{m}, 0\} \right] dG$$
$$+ m(1 - \gamma) + \tau_B + \tau_H + \beta \left[\int_0^\infty V_{IM}(m, d|\varepsilon)dG + W_{IM}(d) \right].$$

Finally, using Equation (1) and (2) together with the fact that $\mathcal{W} = \int_0^\infty V_{IM}(m, d|\varepsilon)dG + W_{IM}(d)$, we arrive at:

$$(1 - \beta)\mathcal{W} = \int_0^\infty \left[\varepsilon^{1/\alpha} f(k_\varepsilon) - k_\varepsilon \right] dG.$$

∎

Proof of Lemma 3.   We first derive the investment schedule for banks. Consider a bank for which the borrowing constraint (10) is slack, so that $\lambda_\varepsilon = 0$. From first-order condition (12) and Equation (8), we find that for this bank

$$k_\varepsilon = \varepsilon \left(\frac{\rho_m}{\rho_d} \right)^\alpha. \tag{35}$$

To ensure that the borrowing constraint (10) is indeed slack, we need $pk_\varepsilon \leq m + \theta pk_s$. Again using Equation (8), it follows that banks face a slack borrowing constraint if and only if

$$\varepsilon \leq \left(\frac{\rho_d}{\rho_m}\right)^\alpha \left(\frac{m}{\rho_d} + \theta k_s\right). \tag{36}$$

Investment by banks for which borrowing constraint (10) is tight, is given by $pk_\varepsilon = m + \theta pk_s$. Again using Equation (8), it follows that for these banks

$$k_\varepsilon = \frac{m}{\rho_d} + \theta k_s \tag{37}$$

Combing Equations (35), (36), and (37), it follows that banks' investment schedule is given by

$$k_\varepsilon = \begin{cases} \varepsilon \left(\rho_m/\rho_d\right)^\alpha & \text{if } \varepsilon \leq \varepsilon' \\ \varepsilon' \left(\rho_m/\rho_d\right)^\alpha & \text{if } \varepsilon > \varepsilon' \end{cases}, \quad \text{where} \quad \varepsilon' = \left(\frac{\rho_d}{\rho_m}\right)^\alpha \left(\frac{m}{\rho_d} + \theta k_s\right). \tag{38}$$

Then, using Equation (38) together with $\int_0^\infty k_\varepsilon \mathrm{d}G = k_s$, gives

$$0 = \varepsilon' - \left(\frac{\rho_d}{\rho_m}\right)^\alpha \frac{m}{\rho_d} - \theta \left[\int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \int_{\varepsilon'}^\infty \varepsilon' \mathrm{d}G\right]. \tag{39}$$

The partial derivative of the RHS of Equation (39) w.r.t. ε' is given by $1 - \theta[1 - G(\varepsilon')]$. With $\theta \in [0, 1]$, the RHS is therefore monotonically increasing in ε'. Moreover, the RHS is negative when $\varepsilon' = 0$ and approaches ∞ when ε' approaches ∞. Hence, Equation (39) uniquely pins down ε'. ∎

Proof of Lemma 4. Equation (13) cannot hold if $i_m < i_n \leq i_p$. Hence, we can focus on cases in which $i_m \geq i_n$ as otherwise demand for reserves carried out of the IM becomes infinitely large.

When $i_n \leq i_m < i_p$, Equation (13) holds if and only if $m_\varepsilon \geq \max\{\bar{m} - pk_s, 0\}$. When $i_n \leq i_m = i_p$, (13) holds if and only if $m_\varepsilon \geq 0$. Finally when $i_n \leq i_p < i_m$, (13) is trivially satisfied.

Similarly, when $i_m = i_n \leq i_p$, Equation (14) is trivially satisfied. When $i_n < i_m \leq i_p$, (14) holds if and only if $0 \leq m_\varepsilon \leq \max\{\bar{m} - pk_s, 0\}$. Finally, when $i_n \leq i_p < i_m$, (14) holds if and only if $m_\varepsilon = 0$.

Combining the insights above, we obtain that m_ε solves (13) and (14) if and only if:

$$m_\varepsilon \leq \begin{cases} \max\{\bar{m} - pk_s, 0\} & \text{if } i_n < i_m \leq i_p \\ 0 & \text{if } i_m > i_p \geq i_n \end{cases}, \quad m_\varepsilon \geq \begin{cases} \max\{\bar{m} - pk_s, 0\} & \text{if } i_n \leq i_m < i_p \\ 0 & \text{if } i_m \geq i_p \geq i_n \end{cases}.$$

Using Equation (8) and defining $\bar{m}' = \max\{\bar{m} - \rho_d k_s, 0\}$, gives Equation (19). ∎

Proof of Proposition 5. We distinguish between money market equilibria without and with excess reserves: $\int_0^\infty m_\varepsilon dG = 0$ and $\int_0^\infty m_\varepsilon dG > 0$, respectively.

In an equilibrium without excess reserves, clearance of the money market implies $m = \int_0^\infty pk_\varepsilon dG$. Using Equations (8) and (16), the latter becomes

$$\frac{m}{\rho_d} = \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right]. \tag{40}$$

Using Equation (15) in Equation (40), we find that

$$\varepsilon' = (1 + \theta) \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right]. \tag{41}$$

With $\theta = 0$, Equation (41) holds only if $\varepsilon' = 0$. According to Equation (40), that must however imply $m = 0$, contradicting the notion of an equilibrium in Definition 1. With $\theta > 0$, there is a unique strictly positive value for ε' that satisfies Equation (41), which we have defined as $\underline{\varepsilon}$. It follows that the money market rate satisfies

$$\frac{m}{\rho_d} = \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\underline{\varepsilon}} \varepsilon dG + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon} dG\right]. \tag{42}$$

It remains to check whether $\int_0^\infty m_\varepsilon dG = 0$, which must imply $m_\varepsilon = 0$ for all ε, is in line with Equation (19). By construction $\bar{m}' = \min\{\bar{m} - m, 0\}$. Clearly, with $\bar{m} \leq m$, $m_\varepsilon = 0$ requires $\rho_m \leq \rho_n$ according to Equation (19). Also, with $\bar{m} > m$, $m_\varepsilon = 0$ requires $\rho_m \leq \rho_p$ according to Equation (19). Hence, in an equilibrium without excess reserves

$$\rho_m \leq \begin{cases} \rho_n & \text{if } \bar{m} \leq m \\ \rho_p & \text{if } \bar{m} > m \end{cases}. \tag{43}$$

In an equilibrium with excess reserves, $m_\varepsilon > 0$ for at least some ε and $pk_s < m$. When $\bar{m} < m$, then $m_\varepsilon > \bar{m}'$ for at least some ε. Otherwise $m = pk_s + \int_0^\infty m_\varepsilon dG \leq pk_s + \max\{\bar{m} - pk_s, 0\}$, which implies either $pk_s \geq m$ or $\bar{m} \geq m$; a contradiction. Hence, Equation (19) holds if and only if $\rho_m = \rho_n$. Similarly, when $\bar{m} > m$, then $m_\varepsilon < \bar{m}'$ for at

least some ε. First, because $pk_s < m$ and $\bar{m} > m$, we have $\bar{m}' > 0$. Second, if $m_\varepsilon \geq \bar{m}'$ for all ε, then money market clearance implies that $m = \int_0^\infty [pk_\varepsilon + m_\varepsilon]dG \geq pk_s + \bar{m}' = \bar{m}$; a contradiction. Hence, because $m_\varepsilon > 0$ for at least some ε and $m_\varepsilon < \bar{m}'$ for some ε, Equation (19) requires that $\rho_m = \rho_p$. Finally, with $m = \bar{m}$, we must have either $\rho_p = \rho_n$ or $m_\varepsilon = \bar{m}'$. Clearly, with $\rho_p < \rho_n$ there cannot be banks with $m_\varepsilon > \bar{m}'$ as well as banks with $m_\varepsilon < \bar{m}'$. Next, because $pk_s < m$ and $\bar{m} = m$, we again have $\bar{m}' > 0$. Moreover, when $m_\varepsilon < \bar{m}'$ for all ε then $m = \int_0^\infty [pk_\varepsilon + m_\varepsilon]dG < pk_s + \bar{m}' = \bar{m}$ and when $m_\varepsilon > \bar{m}'$ for all ε then $m = \int_0^\infty [pk_\varepsilon + m_\varepsilon]dG > pk_s + \bar{m}' = \bar{m}$; both contradictions. With $m_\varepsilon > 0$ for at least some ε and either $\rho_p = \rho_n$ or $m_\varepsilon = \bar{m}' > 0$, it follows that Equation (19) is satisfied for all $\rho_m \in [\rho_p, \rho_n]$. In an equilibrium with excess reserves we therefore find

$$\rho_m \in \begin{cases} \{\rho_n\} & \text{if } \bar{m} < m \\ [\rho_p, \rho_n] & \text{if } \bar{m} = m \\ \{\rho_p\} & \text{if } \bar{m} > m \end{cases} . \tag{44}$$

To ensure that the money market clears with $\int_0^\infty m_\varepsilon dG > 0$, we need $m > \int_0^\infty pk_\varepsilon dG$. Using Equation (16), the latter becomes

$$\frac{m}{\rho_d} > \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG \right] . \tag{45}$$

With $\theta = 0$, Equation (15) implies that Equation (45) is satisfied when $m > 0$. With $\theta > 0$, Equation (15) implies that Equation (45) is satisfied when $\varepsilon' > \underline{\varepsilon}$, which in turn is satisfied whenever

$$\frac{m}{\rho_d} > \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\underline{\varepsilon}} \varepsilon dG + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon} dG \right] .$$

Because $\lim_{\theta \to 0} \underline{\varepsilon} = 0$, the model exhibits no discontinuity at $\theta = 0$. Without loss, matters in the money market are therefore characterized by a floor on the money market rate given by Equation (43) and the following equilibrium condition

$$\frac{m}{\rho_d} \geq \left(\frac{\rho_m}{\rho_d}\right)^\alpha \left[\int_0^{\underline{\varepsilon}} \varepsilon dG + \int_{\underline{\varepsilon}}^\infty \underline{\varepsilon} dG \right], \quad \text{with} = \text{if} \quad \rho_m < \begin{cases} \rho_n & \text{if } \bar{m} < m \\ \rho_p & \text{if } \bar{m} \geq m \end{cases} .$$

∎

Proof of Proposition 6. To derive Equation (23), differentiate $V_{IM}(m, d|\varepsilon)$ with respect to m to get

$$V_{IM}^m(m, d|\varepsilon) = 1/\rho_m + \lambda_\varepsilon.$$

Using Equation (12) to replace λ_ε and using Equation (8) to replace p, yields

$$V_{IM}^m(m,d|\varepsilon) = \frac{\varepsilon^{1/\alpha} f'(k_\varepsilon)}{\rho_d}. \tag{46}$$

Then, use Equation (46) in Equation (6), which should hold with equality when $m > 0$, to get

$$\frac{\gamma}{\beta} = \int_0^\infty \frac{\varepsilon^{1/\alpha} f'(k_\varepsilon)}{\rho_d} dG.$$

Finally, recall that $f'(k) = k^{-1/\alpha}$ and replace the quantities k_ε using Lemma 3 to get

$$\frac{\rho_m \gamma}{\beta} = \int_0^{\varepsilon'} dG + \int_{\varepsilon'}^\infty (\varepsilon/\varepsilon')^{1/\alpha} dG, \tag{47}$$

which is the same as Equation (23), because $\rho_m = \rho_n$ in an equilibrium with full pass-through of NIR to money market rates. With $\bar{m} < m$, combining Equation (20) and our definition of ε' in Equation (15), Proposition 5 implies that $\rho_m = \rho_n$ if $\varepsilon' \geq \underline{\varepsilon}$. The RHS of (23) is strictly decreasing in ε', approaches 1 if $\varepsilon' \to \infty$, and approaches ∞ when $\varepsilon' \to 0$. It follows that an ε' that solves Equation (23) exists if and only if $\gamma \rho_n \geq \beta$ and satisfies $\varepsilon' \geq \underline{\varepsilon}$ if and only if $\gamma \rho_n / \beta \leq \int_0^{\underline{\varepsilon}} dG + \int_{\underline{\varepsilon}}^\infty (\varepsilon/\underline{\varepsilon})^{1/\alpha} dG$. ∎

Proof of Lemma 7. To derive the critical value $\tilde{\varepsilon}$, we consider the meaningful case in which $i_n < i_p$. All banks that are constrained, i.e. $\varepsilon > \varepsilon'$, have $\rho_d k_\varepsilon = m + \theta \rho_d k_s$ and $m_\varepsilon \geq \max\{\bar{m} - \rho_d k_s, 0\}$. Hence, net borrowing in the money market for these banks satisfies

$$z_\varepsilon \geq \theta \rho_d k_s + \max\{\bar{m} - \rho_d k_s, 0\}, \quad \text{for all } \varepsilon > \varepsilon'.$$

Therefore, $z_\varepsilon > 0$ for all $\varepsilon > \varepsilon'$ if $\theta > 0$ and if $\theta = 0$, $z_\varepsilon > 0$ for all $\varepsilon > \varepsilon'$ if $\max\{\bar{m} - \rho_d k_s, 0\} > 0$.

Banks that are unconstrained, i.e. $\varepsilon \leq \varepsilon'$, have $k_\varepsilon = \varepsilon (\rho_m/\rho_d)^\alpha$. Since k_ε is increasing in ε, there exists a critical value $\tilde{\varepsilon}$ such that

$$\rho_d k_{\tilde{\varepsilon}} + \max\{\bar{m} - \rho_d k_s, 0\} = m. \tag{48}$$

For banks with $\varepsilon < \tilde{\varepsilon}$, $z_\varepsilon < 0$ is feasible. We can rewrite Equation (48) as follows

$$k_{\tilde{\varepsilon}} = \frac{m}{\rho_d} - \max\left\{\frac{\bar{m}}{m}\frac{m}{\rho_d} - k_s, 0\right\}.$$

Then, use $k_{\tilde\varepsilon} = \tilde\varepsilon \left(\rho_m/\rho_d\right)^\alpha$ and Equation (15) to get

$$\tilde\varepsilon = \varepsilon' - \theta \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG \right]$$
$$- \max\left\{ \frac{\bar m}{m}\left[\varepsilon' - \theta\left(\int_0^{\varepsilon'}\varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right)\right] - (\rho_d/\rho_m)^\alpha k_s, 0\right\}. \quad (49)$$

Using that $k_s = \int_0^\infty k_\varepsilon dG$ in combination with Equation (16), implies that Equation (49) becomes

$$\tilde\varepsilon = \varepsilon' - \theta\left[\int_0^{\varepsilon'}\varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right]$$
$$- \max\left\{\varepsilon'\frac{\bar m}{m} - \left(1 + \theta\frac{\bar m}{m}\right)\left[\int_0^{\varepsilon'}\varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right], 0\right\},$$

where $\varepsilon' \leq \tilde\varepsilon$ is automatically satisfied. Also, $\varepsilon' < \tilde\varepsilon$ when $\theta > 0$.

Clearly, with $\theta > 0$ it follows that all banks with $\varepsilon > \tilde\varepsilon$ borrow in the money market. Without loss we can assume that all banks with $\varepsilon < \tilde\varepsilon$ then lend in the money market.

With $\theta = 0$, if $\max\{\bar m - \rho_d k_s, 0\} > 0$ all banks with $\varepsilon > \tilde\varepsilon$ borrow in the money market. Without loss we can assume again that all banks with $\varepsilon < \tilde\varepsilon$ then lend in the money market. If $\max\{\bar m - \rho_d k_s, 0\} = 0$, which with $\theta = 0$ holds if and only if $\tilde\varepsilon = \varepsilon'$, we can assume without loss that banks are not active in the money market; $z_\varepsilon = 0$ for all ε.

To derive the critical value $\hat\varepsilon$, note that $k_\varepsilon^* = \varepsilon$ and that $k_\varepsilon \leq \varepsilon(\rho_m/\rho_d)^\alpha$. Thus, ε for which $k_\varepsilon > \varepsilon$ exist only if $1 < (\rho_m/\rho_d)^\alpha$. When $1 < (\rho_m/\rho_d)^\alpha$, it follows from Equation (16) that $k_\varepsilon > \varepsilon$ if and only if $\varepsilon < \varepsilon'(\rho_m/\rho_d)^\alpha$. It follows that $\hat\varepsilon = \varepsilon'(\rho_m/\rho_d)^\alpha$. Using that $\rho_m = \rho_n$ with full pass-through of NIR to money market rates, gives Equation (27). ∎

Proof of Proposition 8. To prove the Proposition, define $\Phi_0 \equiv \rho_n\gamma/\beta$, $\Phi_1 \equiv \int_0^{\varepsilon'} dG$, and $\Phi_2 \equiv \int_{\varepsilon'}^\infty (\varepsilon/\varepsilon')^{1/\alpha} dG$. First, we derive $\frac{d\varepsilon'}{d\rho_n}$. From Equation (23), ε' satisfies

$$\frac{\rho_n\gamma}{\beta} = \int_0^{\varepsilon'} dG + \int_{\varepsilon'}^\infty (\varepsilon/\varepsilon')^{1/\alpha} dG.$$

The derivative satisfies

$$\frac{d\varepsilon'}{d\rho_n} = -\alpha\varepsilon'(\gamma/\beta)\Phi_2^{-1} = -\alpha\varepsilon'\Phi_0\Phi_2^{-1}\rho_n^{-1}.$$

Rearranging yields

$$\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n}\frac{\rho_n}{\varepsilon'} = -\alpha\Phi_0\Phi_2^{-1} < 0,$$

indicating that there are more constrained banks.

Next, from Equation (26)

$$\frac{\mathrm{d}\tilde{\varepsilon}}{\mathrm{d}\rho_n} = \left\{1 - \theta[1 - G(\varepsilon')] - \mathcal{I}\left[\frac{\bar{m}}{m} - \left(1 + \theta\frac{\bar{m}}{m}\right)[1 - G(\varepsilon')]\right]\right\}\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n}, \tag{50}$$

where

$$\mathcal{I} = \begin{cases} 0 & \text{if } \varepsilon\bar{m}/m - (1 + \theta\bar{m}/m)[\int_0^{\varepsilon'}\varepsilon\mathrm{d}G + \int_{\varepsilon'}^{\infty}\varepsilon'\mathrm{d}G] < 0 \\ 1 & \text{if } \varepsilon\bar{m}/m - (1 + \theta\bar{m}/m)[\int_0^{\varepsilon'}\varepsilon\mathrm{d}G + \int_{\varepsilon'}^{\infty}\varepsilon'\mathrm{d}G] \geq 0 \end{cases}.$$

Rearrange Equation (50) as

$$\frac{\mathrm{d}\tilde{\varepsilon}}{\mathrm{d}\rho_n} = \left\{\left(1 - \mathcal{I}\frac{\bar{m}}{m}\right)\left(1 - \theta[1 - G(\varepsilon')]\right) + \mathcal{I}[1 - G(\varepsilon')]\right\}\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n}.$$

With $\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n} < 0$, $\theta \in [0,1]$, and $\bar{m} < m$ we clearly have $\frac{\mathrm{d}\tilde{\varepsilon}}{\mathrm{d}\rho_n} < 0$, indicating that there are more banks that borrow in the money market.

Then, from Equation (27) and using that $\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n}\frac{\rho_n}{\varepsilon'} = -\alpha\Phi_0\Phi_2^{-1}$, we obtain

$$\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n}\frac{\rho_n}{\hat{\varepsilon}} = \alpha\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n}\frac{\rho_n}{\rho_d} - \Phi_0\Phi_2^{-1}\right).$$

With perfect transmission, $\rho_n = \rho_d$ and $\mathrm{d}\rho_n = \mathrm{d}\rho_p$ so $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$. With imperfect transmission, $\mathrm{d}\rho_d = 0$. Using that $\Phi_0 = \Phi_1 + \Phi_2$, it also follows that $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$. So, when $\varepsilon' < \hat{\varepsilon}$ less banks overinvest.

Regarding investment quantities, using Lemma (3) and $\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n}\frac{\rho_n}{\varepsilon'} = -\alpha\Phi_0\Phi_2^{-1}$, we have

$$\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n}\frac{\rho_n}{k_\varepsilon} = \begin{cases} \alpha\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n}\frac{\rho_n}{\rho_d}\right) & \text{if } \varepsilon < \varepsilon' \\ \alpha\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n}\frac{\rho_n}{\rho_d} - \Phi_0\Phi_2^{-1}\right) & \text{if } \varepsilon > \varepsilon' \end{cases}.$$

Clearly, for $\varepsilon < \varepsilon'$ with perfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} = 0$ and with imperfect transmission $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} > 0$. Moreover, for $\varepsilon > \varepsilon'$ we have $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n}\frac{\rho_n}{k_\varepsilon} = \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n}\frac{\rho_n}{\hat{\varepsilon}}$ so $\frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} < 0$ with both perfect and imperfect transmission. Hence, constrained banks invest less. Unconstrained banks invest more with imperfect transmission and their investment remains unaffected with perfect transmission. ∎

Proof of Proposition 9. First, we derive the expression for $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n}$. Take the

derivative of Equation (2) with respect to ρ_n to get

$$\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = \int_0^{\varepsilon'} \left[\varepsilon^{(1/\alpha)} f'(k_\varepsilon) - 1\right] \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} dG + \int_{\varepsilon'}^{\infty} \left[\varepsilon^{(1/\alpha)} f'(k_{\varepsilon'}) - 1\right] \frac{\mathrm{d}k_{\varepsilon'}}{\mathrm{d}\rho_n} dG.$$

Note that ε' depends on ρ_n but the two changes in the integral bounds cancel each other out. We can use Proposition 8 to write this expression as follows

$$\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = \int_0^{\varepsilon'} \left[\varepsilon^{(1/\alpha)} (k_\varepsilon)^{-1/\alpha} - 1\right] \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} dG + \int_{\varepsilon'}^{\infty} \left[\varepsilon^{(1/\alpha)} (k_{\varepsilon'})^{-1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG.$$

Now use Lemma 3 to replace $k_\varepsilon = \varepsilon (\rho_n/\rho_d)^\alpha$ and $k_{\varepsilon'} = \varepsilon' (\rho_n/\rho_d)^\alpha$ to get

$$\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = \int_0^{\varepsilon'} [\rho_d/\rho_n - 1] \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} dG + \int_{\varepsilon'}^{\infty} \left[(\rho_d/\rho_n)(\varepsilon/\varepsilon')^{1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG.$$

Thus, the derivative $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n}$ can be broken into three terms as follows:

$$\frac{d(1-\beta)\mathcal{W}}{d\rho_n} = A + B + C, \text{ where}$$

$$A \equiv \int_0^{\varepsilon'} [\rho_d/\rho_n - 1] \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} dG,$$

$$B \equiv \int_{\varepsilon'}^{\hat{\varepsilon}} \left[(\rho_d/\rho_n)(\varepsilon/\varepsilon')^{1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG,$$

$$C \equiv \int_{\hat{\varepsilon}}^{\infty} \left[(\rho_d/\rho_n)(\varepsilon/\varepsilon')^{1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG.$$

Using $(k_{\varepsilon'})^{-1/\alpha} = (1/\varepsilon')^{1/\alpha} (\rho_d/\rho_n)$ and rearranging $\left[(\rho_d/\rho_n)(\varepsilon/\varepsilon')^{1/\alpha} - 1\right] = \left[(\varepsilon/k_{\varepsilon'})^{1/\alpha} - 1\right]$ we obtain

$$\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = A + B + C, \text{ where}$$

$$A \equiv \rho_d \int_0^{\varepsilon'} (i_n - i_d) \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} dG,$$

$$B \equiv \int_{\varepsilon'}^{\hat{\varepsilon}} \left[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG$$

$$C \equiv \int_{\hat{\varepsilon}}^{\infty} \left[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1\right] \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} dG.$$

Now, we prove $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} < 0$. We restrict attention to a case in which $\gamma\rho_n > \beta$, as $\varepsilon' = \infty$ otherwise. First, we suppose that $\rho_n > \rho_d$ (NIR case). Here, because $A \leq 0$ and $\frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} < 0$, it suffices to show that $\int_{\varepsilon'}^{\infty}[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1]dG > 0$. Use that $(\hat{\varepsilon}/\varepsilon')^{1/\alpha} = \rho_n/\rho_d$

and that Equation (23) implies $(\gamma \rho_n - \beta)/\beta = \int_{\varepsilon'}^{\infty} [(\varepsilon/\varepsilon')^{1/\alpha} - 1] \mathrm{d}G$, to find

$$
\begin{aligned}
\int_{\varepsilon'}^{\infty} [(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1] \mathrm{d}G &= \frac{\rho_d}{\rho_n} \left[\frac{\gamma \rho_n - \beta}{\beta} + \frac{\rho_d - \rho_n}{\rho_d} \left[1 - G(\varepsilon') \right] \right] \\
&\geq \frac{\rho_d}{\rho_n} \left[\frac{\beta \rho_n / \rho_d - \beta}{\beta} + \frac{\rho_d - \rho_n}{\rho_d} \left[1 - G(\varepsilon') \right] \right] \\
&= \frac{\rho_n - \rho_d}{\rho_n} G(\varepsilon') > 0,
\end{aligned}
$$

where the second step uses $\gamma \rho_d \geq \beta$ and the third step uses $\rho_n > \rho_d$.

Next, we drop our supposition $\rho_n > \rho_d$ and note that $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} = \Omega_1 - \Omega_2$, with

$$
\Omega_1 = \int_0^{\infty} \left(\frac{\varepsilon}{k_\varepsilon} \right)^{1/\alpha} \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} \mathrm{d}G \quad \text{and} \quad \Omega_2 = \int_0^{\infty} \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} \mathrm{d}G.
$$

Using Lemma (3) together with insights from the proof of Proposition 8, with imperfect transmission we find for Ω_1

$$
\begin{aligned}
\Omega_1 &= \alpha \frac{1}{\rho_n} \left(\frac{\rho_n}{\rho_d} \right)^{\alpha-1} \left[\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right) \int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \varepsilon' \left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} - \Phi_0 \Phi_2^{-1} \right) \int_{\varepsilon'}^{\infty} \left(\frac{\varepsilon}{\varepsilon'} \right)^{1/\alpha} \mathrm{d}G \right] \\
&= \alpha \frac{1}{\rho_n} \left(\frac{\rho_n}{\rho_d} \right)^{\alpha-1} \left[\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right) \int_0^{\varepsilon'} (\varepsilon - \varepsilon') \mathrm{d}G - \varepsilon' \Phi_0 \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right],
\end{aligned}
$$

where the second step uses how Φ_0, Φ_1, and Φ_2 are defined. For Ω_2 we find

$$
\Omega_2 = \alpha \frac{1}{\rho_n} \left(\frac{\rho_n}{\rho_d} \right)^{\alpha} \left[\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right) \int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} - \Phi_0 \Phi_2^{-1} \right) \int_{\varepsilon'}^{\infty} \varepsilon' \mathrm{d}G \right].
$$

With perfect transmission, meaning that $\frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} = 1$, we find

$$
\begin{aligned}
\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} &\propto -\int_{\varepsilon'}^{\infty} \left[\frac{\rho_d}{\rho_n} \left(\frac{\varepsilon}{\varepsilon'} \right)^{1/\alpha} - 1 \right] \mathrm{d}G \\
&= -\int_{\varepsilon'}^{\infty} \left[\left(\frac{\varepsilon}{\hat{\varepsilon}} \right)^{1/\alpha} - 1 \right] \mathrm{d}G
\end{aligned}
$$

With $\rho_n > \rho_d$, we have already established that $\int_{\varepsilon'}^{\infty} \left[(\varepsilon/\hat{\varepsilon})^{1/\alpha} - 1 \right] > 0$. With $\rho_n \leq \rho_d$, we have $\varepsilon' \geq \hat{\varepsilon}$. It follows that with perfect transmission, $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} < 0$.

With imperfect transmission, meaning that $\frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} = 0$, we find

$$
\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} \propto \frac{\rho_d}{\rho_n} \int_0^{\varepsilon'} (\varepsilon - \varepsilon') \mathrm{d}G - \left[\int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \varepsilon'(1 - \Phi_0 \Phi_2^{-1}) \int_{\varepsilon'}^{\infty} \mathrm{d}G \right]
$$

Note that with imperfect transmission, ρ_d does not depend on ρ_n. Moreover ε' depends only on $\gamma \rho_n / \beta$ and G, and not on ρ_d. As $\int_0^{\varepsilon'} (\varepsilon - \varepsilon') \mathrm{d}G < 0$, the RHS of the equation above must be decreasing in ρ_d. To conclude, evaluate matters when $\rho_d = \rho_n$:

$$\left. \frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} \right|_{\rho_d = \rho_n} \propto \varepsilon' \left(\Phi_0 \Phi_2^{-1} \int_{\varepsilon'}^{\infty} \mathrm{d}G - 1 \right) = \varepsilon' \left(\frac{\Phi_0 [1 - G(\varepsilon')]}{\Phi_0 - G(\varepsilon')} - 1 \right) < 0,$$

where we use that $\Phi_2 = \Phi_0 - G(\varepsilon')$ and $\Phi_0 > 1$ when $\gamma \rho_n > \beta$. It follows that $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} < 0$ when $\rho_n \leq \rho_d$. ∎

Proof of Proposition 10. In equilibrium, aggregate output satisfies

$$Q = \int_0^{\varepsilon'} \varepsilon^{1/\alpha} f(k_\varepsilon) \mathrm{d}G + \int_{\varepsilon'}^{\infty} \varepsilon^{1/\alpha} f(k_{\varepsilon'}) \mathrm{d}G.$$

The total derivative of Q satisfies

$$\frac{\mathrm{d}Q}{\mathrm{d}\rho_n} = \int_0^{\varepsilon'} \varepsilon^{1/\alpha} f'(k_\varepsilon) \frac{\mathrm{d}k_\varepsilon}{\mathrm{d}\rho_n} \mathrm{d}G + \int_{\varepsilon'}^{\infty} \varepsilon^{1/\alpha} f'(k_{\varepsilon'}) \frac{\mathrm{d}\hat{\varepsilon}}{\mathrm{d}\rho_n} \mathrm{d}G.$$

Using the results of the proof of Proposition 8, this equation can be rearranged to

$$\frac{\mathrm{d}Q}{\mathrm{d}\rho_n} = \alpha \frac{1}{\rho_n} \left(\frac{\rho_n}{\rho_d} \right)^{\alpha-1} \left[\left(1 - \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right) \int_0^{\varepsilon'} (\varepsilon - \varepsilon') \mathrm{d}G - \varepsilon' \Phi_0 \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} \right].$$

For imperfect transmission $\frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} = 0$ and for perfect transmission $\frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} = 1$. Clearly, $\frac{\mathrm{d}Q}{\mathrm{d}\rho_n} < 0$ for both cases. Thus, a decrease in i_n has a negative effect on aggregate output. ∎

Proof of Proposition 11. To derive the effect of an increase in ρ_n on the real value of reserves, we take the total derivative of Equation (15).

$$\mathrm{d}\varepsilon' = \frac{\mathrm{d}m}{\rho_d} \left(\frac{\rho_d}{\rho_n} \right)^\alpha - \frac{m}{\rho_d} \left(\frac{\rho_d}{\rho_n} \right)^\alpha \frac{\mathrm{d}\rho_d}{\rho_d} + \theta \left[\int_{\varepsilon'}^{\infty} \mathrm{d}\varepsilon' dG \right] + \alpha \frac{m}{\rho_d} \left(\frac{\rho_d}{\rho_n} \right)^\alpha \frac{\mathrm{d}\rho_d}{\rho_d} - \alpha \frac{m}{\rho_d} \left(\frac{\rho_d}{\rho_n} \right)^\alpha \frac{\mathrm{d}\rho_n}{\rho_n}$$

Rearranging yields

$$\frac{\mathrm{d}m}{\mathrm{d}\rho_n} \frac{\rho_n}{m} = \frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n} \rho_n \frac{\rho_d}{m} \left(\frac{\rho_n}{\rho_d} \right)^\alpha \left[1 - \theta \int_{\varepsilon'}^{\infty} dG \right] + \frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} (1-\alpha) + \alpha$$

The term $\frac{\mathrm{d}\rho_d}{\mathrm{d}\rho_n} \frac{\rho_n}{\rho_d} (1-\alpha)$ is zero for imperfect transmission and negative for perfect transmission since $\alpha > 1$. To show that $\frac{\mathrm{d}m}{\mathrm{d}\rho_n} \frac{\rho_n}{m} < 0$ for both perfect and imperfect transmission

it therefore suffices to show that

$$\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n} \rho_n \frac{\rho_d}{m} \left(\frac{\rho_n}{\rho_d}\right)^\alpha \left[1 - \theta \int_{\varepsilon'}^\infty dG\right] + \alpha < 0. \tag{51}$$

From the proof of Proposition 8,

$$\frac{\mathrm{d}\varepsilon'}{\mathrm{d}\rho_n} = -\alpha\varepsilon' \Phi_0 \Phi_2^{-1} \rho_n^{-1}.$$

Using this in Equation (51) and then rearranging terms yields

$$\alpha \left[1 - \varepsilon' \frac{\Phi_0}{\Phi_2} \frac{\rho_d}{m} \left(\frac{\rho_n}{\rho_d}\right)^\alpha \left[1 - \theta \int_{\varepsilon'}^\infty dG\right]\right] < 0.$$

Thus, it suffices to show that

$$\varepsilon' \frac{\Phi_0}{\Phi_2} \frac{\rho_d}{m} \left(\frac{\rho_n}{\rho_d}\right)^\alpha \left[1 - \theta \int_{\varepsilon'}^\infty dG\right] > 1.$$

Rearranging yields

$$\frac{\Phi_0}{\Phi_2} \frac{\varepsilon' - \theta \int_{\varepsilon'}^\infty \varepsilon' dG}{\frac{m}{\rho_d}\left(\frac{\rho_d}{\rho_n}\right)^\alpha} > 1$$

Using Equation (15) yields

$$\frac{\Phi_0}{\Phi_2} \frac{\varepsilon' - \theta \int_{\varepsilon'}^\infty \varepsilon' dG}{\varepsilon' - \theta \left[\int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^\infty \varepsilon' dG\right]} \geq \frac{\Phi_0}{\Phi_2} \geq 1,$$

with strict inequalities if ε' is bounded, i.e. $\gamma\rho_n > \beta$. ∎

Proposition 15 *Real interest payments of banks, denoted $\mathcal{P}$ are defined as*

$$\mathcal{P} = i_p \bar{m} + i_n (m - \bar{m}). \tag{52}$$

Proof of Proposition 15. Banks receive interest payments $\mathcal{P}$ on their reserve holdings from the central bank

$$\mathcal{P} = \int_0^\infty [i_p \min\{\hat{m}_\varepsilon, \bar{m}\}]\,\mathrm{d}G + \int_0^\infty [i_n \max\{\hat{m}_\varepsilon - \bar{m}, 0\}]\,\mathrm{d}G.$$

First, consider the case $\bar{m} - \rho_d k_s > 0$. Recall, $\hat{m}_\varepsilon = m_\varepsilon + pk_s$. Using Lemma 4 and Proposition 5, $m_\varepsilon = \bar{m} - \rho_d k_s$ for $\varepsilon > \tilde{\varepsilon}$ and $m_\varepsilon > \bar{m} - \rho_d k_s$ for $\varepsilon \leq \tilde{\varepsilon}$. Rearranging

yields

$$\mathcal{P} = \int_0^\infty \bar{m}\mathrm{d}G + \int_0^{\tilde{\varepsilon}} i_n(m_\varepsilon + \rho_d k_s - \bar{m})\mathrm{d}G.$$

Using the market clearing condition in the IM market, $m = \int_0^\infty m_\varepsilon \mathrm{d}G + \rho_d k_s$ and rearranging yields

$$\int_0^{\tilde{\varepsilon}} m_\varepsilon \mathrm{d}G = m - \rho_d k_s - \int_{\tilde{\varepsilon}}^\infty \bar{m}\mathrm{d}G + \int_{\tilde{\varepsilon}}^\infty \rho_d k_s \mathrm{d}G.$$

Plugging this into the expression for banks' real interest payments and rearranging yields

$$\mathcal{P} = i_p \bar{m} + i_n(m - \bar{m}).$$

Second, consider the case $\bar{m} - \rho_d k_s \leq 0$. Since $m_\varepsilon \geq 0$ by Equation (9), we have that $\hat{m}_\varepsilon \geq \bar{m}$ for all ε. Therefore,

$$\mathcal{P} = i_p \bar{m} + i_n \left(\int_0^\infty \hat{m}_\varepsilon \mathrm{d}G - \bar{m} \right).$$

Using that $\hat{m}_\varepsilon = m_\varepsilon + pk_s$ and that money market clearance implies $m = \int_0^\infty m_\varepsilon \mathrm{d}G + pk_s$, yields

$$\mathcal{P} = i_p \bar{m} + i_n(m - \bar{m}).$$

∎

Proof of Proposition 12. From the proof of Proposition 8, we know $\frac{\mathrm{d}(1-\beta)\mathcal{W}}{\mathrm{d}\rho_n} < 0$ for $\rho_n \gamma > \beta$. It follows from continuity that the Friedman rule is optimal. Under the Friedman rule, the LHS of Equation (23) equals 1. The RHS of Equation (23) can only equal 1 if $\varepsilon' \to \infty$. Thus, the Friedman rule implies $\varepsilon' \to \infty$.

Note, for $\rho_n = \rho_d$, the Friedman rule implements the first-best allocation, as all banks invest the first-best quantity. After all, $k_\varepsilon = \varepsilon$ for all $\varepsilon < \varepsilon'$ and $\varepsilon' = \infty$. For $\rho_d > \rho_n$, the Friedman rule is optimal, but does not implement the first-best allocation. In this case, all banks underinvest since $k_\varepsilon = \varepsilon(\rho_n/\rho_d)^\alpha < k_\varepsilon^*$. Lastly, if the central bank runs the Friedman rule, the economy cannot be in the NIR case since the Friedman rule would require $\rho_n = \beta/\gamma > \rho_d$, for which an equilibrium does not exist. ∎

Proof of Lemma 13. Suppose that $\rho_n < \rho_m$ in a steady state equilibrium. Equation (11) then implies that banks do not lend out reserves in the money market; $m_\varepsilon + pk_\varepsilon - m \geq 0$ for all ε. Moreover, Definition 1 requires $\int_0^\infty [m - m_\varepsilon - pk_\varepsilon]\mathrm{d}G(\varepsilon) = 0$. It follows that $m - m_\varepsilon - pk_\varepsilon = 0$ for all ε.

Without loss, we can focus on a case in which $\theta + \sigma > 0$, as otherwise borrowing

reserves in during the IM is infeasible and the money market can be ignored. With $\theta + \sigma > 0$ and $m - m_\varepsilon - pk_\varepsilon = 0$ for all ε, the borrowing constraint must be slack for all ε-banks; $\lambda_\varepsilon = 0$ for all ε. But with $\rho_p \leq \rho_n < \rho_m$, Equation (13) then requires $\mu_\varepsilon < 0$ for all ε, which cannot be the case. That means, if banks' borrowing constraints are slack but $\rho_p \leq \rho_n < \rho_m$, banks have an incentive to borrow (additional) reserves in the money market and deposit them at the central bank. ∎

Proof of Proposition 14. First, we need that some banks enter the settlement market with reserves strictly smaller than the exemption threshold. This is the case if and only if $\bar{m} - \rho_d k_s > 0$. Using the latter in Equation (32) and comparing the result with Equation (33), it follows that $\bar{m} - \rho_d k_s > 0$ if $\varepsilon' < \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha$.

Second, combining Equations (6), (30), and (8) with the fact that $V_{IM}^m(m, d|\varepsilon) = (\varepsilon/k_\varepsilon)^{1/\alpha}/p$ yields

$$\frac{\gamma \rho_n}{\beta} = \int_0^{\varepsilon'} dG + \int_{\varepsilon'}^{\varepsilon''} \left(\frac{\varepsilon}{\varepsilon'} \right)^{1/\alpha} dG + \frac{\rho_n/\rho_p - \sigma}{1 - \sigma} \left[\int_{\varepsilon''}^{\varepsilon'''} dG + \int_{\varepsilon'''}^{\infty} \left(\frac{\varepsilon}{\varepsilon'''} \right)^{1/\alpha} dG \right]. \quad (53)$$

Third, combining Equations (32) and (33) with $k_s = \int_0^\infty k_\varepsilon dG$ yields

$$\frac{\varepsilon' - \left[1 - (1-\sigma)\frac{\bar{m}}{m} \right] \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha}{(1-\sigma)\left(1 + \theta\frac{\bar{m}}{m} \right)}$$
$$= \int_0^{\varepsilon'} \varepsilon dG + \int_{\varepsilon'}^{\varepsilon''} \varepsilon' dG + \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha \left[\int_{\varepsilon''}^{\varepsilon'''} \varepsilon dG + \int_{\varepsilon'''}^{\infty} \varepsilon''' dG \right]. \quad (54)$$

Fourth, we show that $\varepsilon' < \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha$ if and only if $\varepsilon' \geq \bar{\varepsilon}$. Let

$$\mathcal{H} = \frac{\varepsilon' - \left[1 - (1-\sigma)\frac{\bar{m}}{m} \right] \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha}{(1-\sigma)\left(1 + \theta\frac{\bar{m}}{m} \right)}$$
$$- \int_0^{\varepsilon'} \varepsilon dG - \int_{\varepsilon'}^{\varepsilon''} \varepsilon' dG - \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha \left[\int_{\varepsilon''}^{\varepsilon'''} \varepsilon dG + \int_{\varepsilon'''}^{\infty} \varepsilon''' dG \right].$$

Observe that $\varepsilon'' = \varepsilon'(\frac{\rho_n/\rho_p - \sigma}{1-\sigma})^\alpha$ pins down ε'' as a function of ε'. Also, $\partial \mathcal{H}/\partial \varepsilon''' < 0$ when $\bar{m} < m$ and $\partial \mathcal{H}/\partial \varepsilon'' \big|_{\varepsilon'' = \varepsilon'(\frac{\rho_n/\rho_p - \sigma}{1-\sigma})^\alpha} = 0$. Therefore, $\varepsilon' < \varepsilon''' \left(\frac{1-\sigma}{\rho_n/\rho_p - \sigma} \right)^\alpha$ holds if and only if

$$\mathcal{H}\Big|_{\varepsilon''' = \varepsilon'\left(\frac{\rho_n/\rho_d - \sigma}{1-\sigma} \right)^\alpha} > 0 \Rightarrow \frac{\bar{m}\varepsilon'}{m + \theta\bar{m}} - \int_0^{\varepsilon'} \varepsilon dG - \int_{\varepsilon'}^{\infty} \varepsilon' dG > 0, \quad (55)$$

where we use that $\varepsilon'' = \varepsilon'''$ if $\varepsilon''' = \varepsilon'\left(\frac{\rho_n/\rho_d - \sigma}{1-\sigma}\right)^\alpha$. In turn, Equation (55) holds if and only if $\varepsilon' > \bar{\varepsilon}$.

Fifth, we show that a pair $(\varepsilon', \varepsilon''')$ that solves both Equations (53) and (54) with $\varepsilon'' = \varepsilon'(\frac{\rho_n/\rho_p - \sigma}{1-\sigma})^\alpha$, and satisfies $\varepsilon' > \bar{\varepsilon}$, is unique. Because $\varepsilon' > \bar{\varepsilon}$, we have that $1 - (m/\bar{m} + \theta)[1 - G(\varepsilon')] > 0$. With $m/\bar{m} + \theta \leq 1$ this is obvious. With $m/\bar{m} + \theta > 1$, note that

$$\mathcal{G} = \varepsilon' - (m/\bar{m} + \theta)\left[\int_0^{\varepsilon'} \varepsilon \mathrm{d}G + \int_{\varepsilon'}^\infty \varepsilon'\right] \tag{56}$$

is first decreasing in ε' (i.e. until ε' is such that $1 - (m/\bar{m} + \theta)[1 - G(\varepsilon')] = 0$) and then increasing in ε'. Since, $\mathcal{G} = 0$ when $\varepsilon' = 0$, it follows that $1 - (m/\bar{m} + \theta)[1 - G(\varepsilon')] > 0$ when $\mathcal{G} > 0$, i.e. when $\varepsilon' > \bar{\varepsilon}$.

With this result, we can perform the following manipulations regarding $\partial \mathcal{H}/\partial \varepsilon'$:

$$\begin{aligned}
\frac{\partial \mathcal{H}}{\partial \varepsilon'} &= \frac{1}{(1-\sigma)(1+\theta\bar{m}/m)} - \int_{\varepsilon'}^{\varepsilon''} \mathrm{d}G \\
&= \frac{1 - (1-\sigma)(1+\theta\bar{m}/m)\left[G(\varepsilon'') - G(\varepsilon')\right]}{(1-\sigma)(1+\theta\bar{m}/m)} \\
&> \frac{1 - (1-\sigma)(1+\theta\bar{m}/m)[1 - G(\varepsilon')]}{(1-\sigma)(1+\theta\bar{m}/m)} \\
&> \frac{1 - \bar{m}(1-\sigma)/m}{(1-\sigma)(1+\theta\bar{m}/m)} \\
&> 0.
\end{aligned}$$

In equilibrium $\mathcal{H} = 0$, so Equation (54) pins down ε''' as a strictly increasing function of ε'. Because the RHS of Equation (53) is strictly decreasing in both ε' and ε''', and does not respond to small changes in ε'' when $\varepsilon'' = \varepsilon'(\frac{\rho_n/\rho_p - \sigma}{1-\sigma})^\alpha$, it follows that an equilibrium is unique.

Sixth, to have existence of a pair $(\varepsilon', \varepsilon''')$ that solves both Equations (53) and (54), and satisfies $\varepsilon' > \bar{\varepsilon}$, it follows that we need

$$1 \leq \frac{\gamma\rho_n}{\beta} < \int_0^{\bar{\varepsilon}} \mathrm{d}G + \int_{\bar{\varepsilon}}^\infty (\varepsilon/\bar{\varepsilon})^{1/\alpha} \mathrm{d}G.$$

Finally, to ensure that we indeed have an equilibrium, it remains to verify that the money market clears: $m \geq \rho_d k_s + \int_0^\infty \underline{m}_\varepsilon \mathrm{d}G$. Because $\underline{m}_\varepsilon \leq \bar{m} - \rho_d k_s$ by Equation (31), we have that $m \geq \rho_d k_s + \int_0^\infty \underline{m}_\varepsilon \mathrm{d}G$ holds for sure if $m \geq \bar{m}$. ∎

References

Kim Abildgren, Nicolaj Albrechtsen, Mark Strøm Kristoffersen, Søren Truels Nielsen, and Rasmus Tommerup. The short-term Danish interbank market before, during and after the financial crisis. *Danmarks Nationalbank Working Papers No. 99*, 2015.

Gara Afonso and Ricardo Lagos. Trade dynamics in the market for federal funds. *Econometrica*, 83(1):263–313, 2015.

Ruchir Agarwal and Miles S. Kimball. Enabling deep negative interest rates to fight recessions: A guide. *IMF Working Paper WP/19/84*, 2019.

Carlo Altavilla, Lorenzo Burlon, Mariassunta Giannetti, and Sarah Holton. Is there a zero lower bound? The effects of negative policy rates on banks and firms. *European Central Bank Working Paper Series No. 2289*, 2019.

Adrien Amzallag, Alessandro Calza, Dimitris Georgarakos, and João Sousa. Monetary policy transmission to mortgages in a negative interest rate environment. *European Central Bank Working Paper No. 2243*, 2019.

David M. Arseneau. How would us banks fare in a negative interest rate environment. *Finance and Economics Discussion Series 2017-030r1, Washingon: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2017.030r1*, 2017.

Bank of Japan. Introduction of "quantitative and qualitative monetary easing with a negative interest rate". *Announcements*, 2016.

Christoph Basten and Mike Mariathasan. How banks respond to negative interest rates: Evidence from the swiss exemption threshold. *CESifo Working Paper Series No. 6901*, 2018.

Morten Bech and Cyril Monnet. A search-based model of the interbank money market and monetary policy implementation. *Journal of Economic Theory*, 164:32–67, 2016.

Morton Bech and Aytek Malkhozov. How have central banks implemented negative interest rates? *BIS Quarterly Review*, (1), 2016.

Aleksander Berentsen and Cyril Monnet. Monetary policy in a channel system. *Journal of Monetary Economics*, 55(6):1067–1080, 2008.

Aleksander Berentsen, Alessandro Marchesiani, and Christopher J. Waller. Floor systems for implementing monetary policy: Some unpleasant fiscal arithmetic. *Review of Economics Dynamics*, 17(3):523–542, 2014.

Aleksander Berentsen, Sébastien Kraenzlin, and Benjamin Müller. Exit strategies for monetary policy. *Journal of Monetary Economics*, 99:20–40, 2018.

Tom Bernhardsen and Kathrine Lund. Negative interest rates: Central bank reserves and liquidity management. *Norges Bank Economic Comments*, No. 2, 2015.

Jef Boeckx, Maite de Sola Perea, and Gert Peersman. The transmission mechanism of credit support policies in the euro area. *European Economic Review*, 124:103403, 2020.

Margherita Bottero, Camelia Minoiu, José-Luis Peydró, Andrea Polo, Andrea Presbitero, and Enrico Sette. Expansionary yet different: Credit supply and real effects of negative interest rate policy. *Bank of Italy Temi Di Discussione (Working Paper) No. 1269*, 2020.

Michael Boutros and Jonathan Witmer. Monetary policy implementation in a negative rate environment. *Journal of Money, Credit and Banking*, 52(2-3):441–470, 2020.

Falk Bräuning and Botao Wu. ECB monetary policy transmission during normal and negative interest rate periods. *Available at SSRN 2940553*, 2017.

Markus K Brunnermeier and Yann Koby. The reversal interest rate: An effective lower bound on monetary policy. *Submitted*, 2019.

Danmarks Nationalbank. Danmarks nationalbank adjusts account conditions for banks and mortgage-credit institutions (current-account limits). *Press Releases*, 2016.

Giovanni Dell'Ariccia, Vikram Haksar, Tommaso Mancini-Griffoli, Kelly Eckhold, Simon Gray, Fei Han, Gee Hee Hong, Erik Lundback, Hiroko Oura, Helene Poirson, et al. Negative interest rate policies-initial experiences and assessments. *IMF Policy Paper*, 2017.

Selva Demiralp, Jens Eisenschmidt, and Thomas Vlassopoulos. Negative interest rates, excess liquidity and retail deposits: Banks' reaction to unconventional monetary policy in the euro area. *European Central Bank Working Paper Series No. 2283*, 2019.

Douglas W Diamond and Philip H Dybvig. Bank runs, deposit insurance, and liquidity. *Journal of Political Economy*, 91(3):401–419, 1983.

Feng Dong and Yi Wen. Optimal monetary policy under negative interest rate. *Federal Reserve Bank of St. Louis Research Division Working Paper Series*, Working Paper 2017-019A, 2017.

ECB. ECB introduces a negative deposit facility interest rate. *Press Release*, 2014.

ECB. ECB introduces two-tier system for remunerating excess liquidity holdings. *Press Release*, 2019.

Gauti B Eggertsson, Ragnar E Juelsrud, Lawrence H Summers, and Ella Getz Wold. Negative nominal interest rates and the bank lending channel. *NBER Working Paper No. 25416*, 2019.

Jens Eisenschmidt and Frank Smets. Negative interest rates: Lessons from the euro area. *Series on Central Banking Analysis and Economic Policies No. 26, Central Banking, Analysis, and Economic Policies Book Series, in Monetary Policy and Financial Stability: Transmission Mechanisms and Policy Implications (edited by lvaro Aguirre & Markus Brunnermeier & Diego Saravia*, 26:013–042, 2019.

Craig H Furfine. Interbank payments and the daily federal funds rate. *Journal of Monetary Economics*, 46(2):535–553, 2000.

Christian Grisse and Silvio Schumacher. The response of long-term yields to negative interest rates: Evidence from Switzerland. *SNB Working Papers 10/2017*, 2017.

Christian Grisse, Signe Krogstrup, and Silvio Schumacher. Lower-bound beliefs and long-term interest rates. *International Journal of Central Banking*, 13(3):165–202, 2017.

Florian Heider, Farzad Saidi, and Glenn Schepens. Life below zero: Bank lending under negative policy rates. *The Review of Financial Studies*, 32(10):3728–3761, 2019.

Harriet Jackson. The international experience with negative policy rates. *Bank of Canada Staff Discussion Paper 2015-13*, 2015.

Carina Moselund Jensen and Morten Spange. Interest rate pass-through and the demand for cash at negative interest rates. *Danmarks Nationalbank Monetary Review, 2nd Quarter*, 2015.

Andreas Jobst and Huidan Lin. Negative interest rate policy (NIRP): Implications for monetary transmission and bank profitability in the Euro area. *IMF Working Papers WP/16/172*, 2016.

Anders Jørgensen and Lars Risbjerg. Negative interest rates. *Danmarks Nationalbank Monetary Review 3rd Quarter Part 1*, 2012.

Wolfgang Lemke and Andreea Vladu. Below the zero lower bound: a shadow-rate term structure model for the Euro area. *European Central Bank Working Paper Series No. 1991*, 2017.

James McAndrews. Negative nominal central bank policy rates: Where is the lower bound? *Speech held at University of Wisconsin*, 2015.

Daniel Orr and W Giles Mellon. Stochastic reserve losses and expansion of bank credit. *The American Economic Review*, 51(4):614–623, 1961.

William Poole. Commercial bank reserve management in a stochastic model: Implications for monetary policy. *The Journal of Finance*, 23(5):769–791, 1968.

David Porcellacchia. Maturity transformation and negative interest rate policies. *Unpublished manuscript*, 2019.

Guillaume Rocheteau, Randall Wright, and Sylvia Xiaolin Xiao. Open market operations. *Journal of Monetary Economics*, 98:114–128, 2018.

Matthew Rognlie. What lower bound? monetary policy with negative interest rates. *Unpublished manuscript, Department of Economics, Harvard University*, 2015.

Tan Schelling and Pascal Towbin. Negative interest rates, deposit funding and bank lending. *SNB Working Papers 5/2020*, 2020.

SNB. Merkblatt zu Negativzinsen auf Girokontoguthaben. 2015.

SNB. Instruction sheet governing negative interest on sight deposit account balances. 2019.

Sveriges Riksbank. Riksbank cuts repo rate to -0.10% per cent, buys government bonds for sek 10 billion and is prepared to more at short notice. *Press Release*, 2015.

Rima Turk. Negative interest rates: How big a challenge for large danish and swedish banks? *IMF Working Paper WP/16/198*, 2016.

Mauricio Ulate. Going negative at the zero lower bound: The effects of negative nominal interest rates. *Federal Reserve Bank of San Francisco Working Paper 2019-21, . https://doi.org/10.24148/wp2019-21*, 2019.

William Whitesell. Interest rate corridors and reserves. *Journal of Monetary Economics*, 53(6):1177–1195, 2006.

Fritz Zurbrügg. Negative interest rates: Necessary from a monetary policy perspective - but with what risks for the banks. *Speech held in German at Volkswirtschaftliche Gesellschaft des Kanton Bern*, 2016.