Bank Competition and Risk Taking in a Zero Interest Environment

Robin Döttling*

University of Amsterdam and Tinbergen Institute

Job Market Paper

November 28, 2017

[Link to the Latest Version]

Abstract

How do near-zero interest rates affect bank competition, risk taking and financial regulation? I study these questions in a tractable dynamic general equilibrium model, in which forward-looking banks compete imperfectly for deposit funding, and deposit insurance may induce excessive risk taking. The zero lower bound (ZLB) distorts bank competition and boosts risk shifting incentives, particularly if rates are expected to remain near-zero for long. At the ZLB, capital regulation becomes a less effective tool to curb moral hazard. When banks cannot pass on the cost of capital to depositors, tight capital requirements erode franchise value, countervailing the usual “skin in the game” effect. Very low interest rates may therefore motivate weaker capital regulation, despite overall higher risk. Complementing existing regulation with policy tools that subsidize the funding cost of banks may improve welfare at the ZLB.

Keywords. Zero lower bound, search for yield, capital regulation, bank competition, risk shifting, franchise value

JEL classifications. G21, G28, E43

*Email: r.j.doettling@uva.nl. Address: Amsterdam Business School, Plantage Muidergracht 12, 1018 TV Amsterdam

I am grateful to Enrico Perotti, David Martinez-Miera, Rafael Repullo, Javier Suárez, Douglas Gale, Frédéric Malherbe, Tanju Yorulmazer, Daisuke Ikeda, Magdalena Rola-Janicka, Germán Gutiérrez, Simas Kucišinas, William Diamond and Thomas Philippon for their valuable comments, feedback and suggestions. Seminar participants at the University of Amsterdam, the Tinbergen Institute, the Bank of England, the CEMFI, and the New York University also provided useful suggestions.
1. Introduction

In December 2008, the Federal Reserve lowered the target range for its main policy rate to near-zero, where it would remain for seven years. In the meantime, the Fed implemented various measures of unconventional monetary policy to push down the longer end of the yield curve. While extremely low interest rates have received considerable attention in monetary economics, the implications of the zero lower bound (ZLB) for banking and financial regulation have been studied less extensively. The goal of this paper is to fill this gap, and its key contribution is to show that the ZLB may undermine the effectiveness of prudential regulation.

The business model of banks is to invest in relatively safe, long-term assets, that are overwhelmingly funded with short-term, liquid liabilities (Hanson et al., 2015). This liquidity transformation allows banks to earn a margin between interest income and expense, supported by their market power in deposit markets (Drechsler et al., 2017a).

This business model comes under pressure when deposit rates are constrained by a (zero) lower bound. The left panel in figure 1 shows the yield on highly rated U.S. corporate bonds, along with the interest expense per unit of deposit funding of the median U.S. bank. Notwithstanding swings in the level of interest rates, the two series move in parallel until 2008. Thereafter, a clear compression in the spread between them is visible (right panel).

Recent contributions highlight that low interest rates may generally weaken bank market power (Drechsler et al., 2016) and stimulate risk taking (e.g. Jiménez et al., 2014; Martinez-Miera and Repullo, 2017). Less emphasis has been put on the distinct effect of the zero lower bound on deposit rates, which is at the core of the analysis here. This focus is motivated by the evidence presented in figure 2, which shows that after 2009 the distribution of deposit rates becomes increasingly right-skewed as rates bunch near zero. Heider et al. (2016) find

---

1 Over the period 1996-2013, the median U.S. bank in the Call Reports data funds 85% of its assets with deposits. While larger banks have easier access to wholesale funding, the median deposit funding is still 76% among banks with balance sheet size greater than $1 billion.

2 Appendix B shows that the spread between the interest income ratio and deposit interest expense ratio follows a similar pattern, and drops around 2007. Using bond yields as a benchmark compares deposit rates to the return on relatively safe assets with a given level of risk. The interest expense ratio is calculated using Call Reports data (series riad4170 divided by rcon2200). Due to the short maturity of deposits it is a good approximation for a bank’s current average deposit rate. The bond yield is the BofA Merrill Lynch US Corporate AAA Effective Yield, retrieved from FRED.
Figure 1: This figure plots the yield on highly rated corporate bonds, along with the interest expense per unit of deposit funding of the median U.S. bank in the call reports data. The right panel plots the spread between the two series.

similar evidence for the Eurozone, suggesting that many banks are unable or unwilling to lower deposit rates into negative territory (also see Bech and Malkhozov, 2016; Eggertsson et al., 2017).

How do banks react to this environment of near-zero interest rates and compressed margins? Does it induce banks to reach for yield, and how should financial regulation react? I study these questions in a tractable dynamic general equilibrium model, in which banks compete imperfectly for deposit funding, and deposit insurance may induce excessive risk taking. Crucially, I allow for bank failure, such that the risk taking decision of forward-looking banks is driven by their franchise value, which in turn is endogenously determined by bank competition (Hellmann et al., 2000).

In the model, firms produce output using capital as the only input and there are two technologies to produce new capital. Households can directly produce new capital, which I interpret as investments in the financial market. Banks can also produce new capital, albeit at a higher cost (perhaps due to the cost of operating a branch network and complying with regulation). The bank’s technology can be interpreted as loans to a bank-dependent sector.

Banks take the return on capital as given and set deposit rates under monopolistic com-

\[3\]While there are some cases of banks charging negative rates, a majority seems hesitant to do so. This seems to be particularly true for retail deposits, which may more easily substitute to cash as a means of payment and storage. Perhaps behavioral biases play a role too, as retail customers may be outraged seeing negative interest rates on their account.
Figure 2: For selected years from 1999-2013, this figure plots the cross-sectional distribution of deposit interest expense ratios across U.S. banks in the Call Reports data. The deposit interest expense ratio is defined as interest expenses per unit of deposits. Figure 10 in the appendix reports histograms for additional years.

petition. Households are willing to accept a lower return on deposits because they carry a liquidity service valued in their utility function. \(^4\)

The representative household’s discount factor alternates according to a two-state Markov process, generating variation in the level of interest rates. A high-rate state represents “normal” times with interest rates well above zero, such as the period 1996-2008 in figure 1. In the low-rate state deposit rates may be constrained by the ZLB, as from 2009-2015.

There is a moral hazard problem, as deposits are insured by the government and shareholders have limited liability. At the same time, banks stand to lose rents upon failure because they have market power. In balance, banks trade off the gains from shifting risk on the deposit

\(^4\)Since equity carries no convenience yield and bank investments are costlier than in the financial market, bank capital is (socially) costly in the model.
insurance against the risk of loss of franchise value.

I use the model to study whether low interest rates induce banks to take more risk - a concern that has first been articulated by Rajan (2005) during the run-up to the financial crisis of 2008. In contrast to other contributions in the search for yield literature (e.g. Acharya and Plantin, 2016; Martinez-Miera and Repullo, 2017), I find that a reduction in interest rates has little effect on bank risk taking, so long as deposit rates are not constrained by the ZLB. The reason is that market power allows banks to pass on reductions in interest rates to depositors and maintain stable interest margins, in line with recent evidence (Drechsler et al., 2017a, also see appendix B for additional evidence). If anything, lower discount rates boost franchise values, inducing banks to take less risk.

In contrast, when deposit rates are constrained by the ZLB, a reduction in lending rates eats into margins and erodes profitability. As a consequence of lower franchise values, banks shift more risk on the deposit insurance, particularly if the yield curve flattens substantially and the ZLB is expected to bind for a long time. Moreover, even in the high-rate state banks take more risk if it is likely that the economy transitions to the low-rate state with a binding ZLB in the future. This dynamic effect highlights that even as the Fed started normalizing rates in 2015, the possibility of falling back to the ZLB in the future may still affect incentives.

In the model, capital regulation limits the leverage banks can take. The optimal level of capital requirements crucially depends on whether the ZLB binds. Calibrating the model to U.S. data and allowing for a state-dependent capital requirement, I find an optimal level around 7-8% in both the high-rate and low-rate state when the ZLB remains slack. In contrast, with a binding ZLB the optimal requirement is lower in the low-rate state, despite overall higher risk taking. Generally, higher capital requirements reduce risk taking incentives as they increase a bank’s “skin in the game”. However, at the ZLB a countervailing effect comes into play that I label “capital overhang”. When banks cannot pass on the cost of capital to depositors, tight capital requirements erode franchise value, with the perverse effect of increasing risk taking incentives. This effect reduces the overall effectiveness of capital regulation in curbing risk shifting incentives, motivating weaker requirements despite overall higher risk taking.

Low profitability in the low-rate state also motivates tighter capital requirements in the

---

5 The exact level depends on how compressed intermediation spreads are.
high-rate state, where the ZLB is still slack. Thus, the model delivers a novel rationale for counter-cyclical capital regulation, in the sense that a binding ZLB motivates lower capital requirements, while the optimal capital requirement is higher when the ZLB is slack.\footnote{The case for counter-cyclical requirements is often made in models with welfare-relevant pecuniary externalities (e.g. Lorenzoni, 2008; Korinek and Simsek, 2016). The argument in the policy debate is that buffers built up in good times should be available to be used in bad times (e.g. Goodhart et al., 2008), and relies on frictions to raising equity. In contrast, the rationale here is relevant even absent any frictions to raising equity and welfare-relevant pecuniary externalities.} The capital overhang effect is also relevant for the debate on whether monetary policy should target financial stability. Since very low interest rates can undermine the effectiveness of prudential policies, it may not be enough to solely rely on macro-prudential policies to ensure financial stability.

I consider as an alternative policy tool a subsidy per unit of deposits, paid to banks whenever the ZLB binds. Such a policy effectively supports interest margins, and resembles funding schemes such as the ECB’s targeted long-term refinancing operations or the Bank of England’s Funding for Lending and Term Funding schemes.\footnote{These policy schemes offer cheap funding to banks, conditional on making loans.} This policy restores incentives as it stabilizes profitability. However, its overall effect on welfare is ambiguous because the taxes raised to fund the subsidy create an additional distortion.

1.1. Related Literature

This paper relates to the literature on “search for yield”, which argues that low interest rates result in an increase in risk taking (Rajan, 2005; Jiménez et al., 2014; DellAriccia et al., 2014; Martinez-Miera and Repullo, 2017; Drechsler et al., 2017b). Closely related, Heider et al. (2016) show in a diff-in-diff setting that negative interest rates in the Eurozone have induced banks that rely relatively more on deposit funding to lend to relatively riskier borrowers.

A priori it is not clear why the level of an equilibrium price (the interest rate) should affect the incentives to take excessive risk. Consequently, most contributions in the literature identify mechanisms that explain how the level of interest rates affects the margin between asset returns and cost of liabilities. Relative to this literature, I have a distinct focus on deposit competition in the presence of a zero lower bound on deposit rates. In normal times, interest margins are determined by market power and hence not affected by the level of...
interest rates. Since additionally lower discount rates boost franchise values, I find that low interest rates may actually decrease risk taking incentives.8

Via the risk taking channel of monetary policy an increase in risk tolerance is an intended effect of low interest rates (Borio and Zhu, 2012; Choi et al., 2016; Acharya and Plantin, 2016). My paper relates to this literature, but focuses on inefficient risk shifting, that is a result of agency problems. The capital overhang effect identified in this paper contributes to the debate on whether monetary policy should target financial stability, as it highlights a relevant interaction between the level of interest rates and the effectiveness of prudential regulation.

My modeling approach is related to Begenau (2016) and Davydiuk (2017), who study optimal capital regulation in quantitative DSGE models with banks. Relative to these papers, a distinct feature of my model is that I allow for market power and bank failure.9

Because I allow for bank failure, risk taking is driven by franchise value, as banks stand to lose rents upon failure (Hellmann et al., 2000). Because I allow for market power, banks have strictly positive franchise values.

For the franchise value effect to be operative I need to explicitly solve the forward-looking problem of banks in the presence of occasionally binding constraints. To simplify this problem, I minimize the number of state variables, implying that I leave out several elements present in Begenau (2016) that would be desirable to include for a quantitative assessment. Therefore, any quantitative results of my paper should be seen as suggestive. Still, I am able to meaningfully calibrate the model and find a realistic optimal level of capital requirements (8%).

That franchise value affects risk taking has been shown in the banking literature (e.g. Hellmann et al., 2000; Perotti and Suarez, 2002; Repullo, 2004; Martinez-Miera and Repullo, 2010). To the best of my knowledge I am the first to explicitly incorporate it in a dynamic general equilibrium model.

Closely related to my paper, Hellmann et al. (2000) argue that capital requirements should be complemented with interest rate ceilings, since capital requirements erode franchise value.

---

8DellAriccia et al. (2014) also show that lower rates may decrease risk taking. If bank leverage is fixed, the overall cost of liabilities of highly levered banks may decrease relative to the return on loans as rates fall. In Begenau (2016) banks have unlimited liability, and in Davydiuk (2017) they are bailed out with probability one.
The ZLB has the opposite effect of an interest rate ceiling, imposing a minimum rate banks have to pay to depositors. In contrast to Hellmann et al. (2000), I find that capital requirements erode franchise value only if the ZLB binds, since otherwise market power allows banks to fully pass on the cost of capital.\textsuperscript{10}

A recent literature studies how monetary policy affects the market power of banks (Drechsler et al., 2016; Scharfstein and Sunderam, 2015) and shadow banks (Xiao, 2017). Drechsler et al. (2016) present the “Deposits Channel” of monetary policy, in which market power allows banks to pass on increases in the Fed Funds rate less than 1-1 to depositors. Bank competition is also at the heart of my model, but I have a distinct focus on the zero lower bound of deposit rates, motivated by the evidence in figure 2. Closely related, Brunnermeier and Koby (2016) introduce the concept of a “reversal rate”, below which monetary policy becomes ineffective. While the authors also highlight the negative effect of low interest rates on bank profitability, my paper has a distinct focus on risk taking and implications for bank capital regulation.

Finally, this paper is related to the macroeconomic literature on the zero lower bound and liquidity traps (e.g. Keynes, 1936; Krugman, 1998; Eggertsson and Woodford, 2003; Eggertsson and Krugman, 2012; Guerrieri and Lorenzoni, 2017). While this literature focuses on monetary and fiscal policy, I show that the ZLB may also constrain the effectiveness of prudential regulation. In fact, I study the effect of the ZLB in a real model, in which interest rates clear the savings-investment market. My economy can therefore be interpreted as one in which the price level and inflation expectations are fixed.

The rest of the paper is organized as follows. Section 2 describes the model setup, equilibrium and calibration. Section 3 characterizes the first best allocation, and section 4 studies the determinants of bank risk taking in the model. Section 5 derives optimal capital regulation, shows how it is affected by the capital overhang effect, and discusses alternative policy options. Finally, section 6 concludes.

\textsuperscript{10}This difference is the result of different modeling choices. The analysis in Hellmann et al. (2000) is based on a Monti-Klein type model of competition. In my paper, the demand for deposits is endogenously derived from household optimization, and banks can pass on the cost of capital 1-1 to depositors.
2. Model Setup

In the model, time runs discretely from $t = 0, \ldots, \infty$. A representative household invests in bank deposits and the financial market, where deposits generate additional utility as they provide liquidity services.

Firms produce output using capital as the only input. There are two technologies to produce new capital, one operated by households, and the other by banks. The former technology represents bank-independent finance, where households directly fund investments through the financial market. Capital produced by banks can be interpreted as bank loans that fund the investment of bank-dependent firms or mortgages. In the remainder I adopt these interpretations, and refer to the technologies as the financial market and bank loans, respectively.\(^{11}\)

Banks compete monopolistically for deposit funding, subject to a zero lower bound on deposit rates.\(^{12}\) Deposits are insured by the government, which funds the insurance through taxes running a balanced budget.\(^{13}\) Moreover, banks choose the riskiness of their loans and are subject to a capital requirement that limits the leverage they can take.

The main focus of the paper is on how the level of interest rates affects risk taking and optimal capital regulation, in particular when the ZLB binds. To generate stochastic variation in interest rates I allow the household’s discount factor to vary according to a 2-state Markov process.

The flow diagram in figure 3 summarizes the timing within a period $t$. In the beginning of period $t$ (stage A) firms produce output and pay households and banks a return on their investments made at $t-1$. Banks use the proceeds to repay depositors and pay a dividend, and firms return their profits to households. Afterwards (stage B), households consume and new investments are made.

\(^{11}\)In the real world, banks lend to firms which in turn make physical investments. Leaving out this extra layer of capital producing firms is equivalent to assuming that there are no frictions between them and banks.

\(^{12}\)Here, the ZLB is exogenously assumed, but it can be motivated by the option to hoard cash with a net return of zero. In this paper, the ZLB plays the role as an off-equilibrium outside option, and its exact motivation is irrelevant.

\(^{13}\)Deposit insurance is taken as a given institutional feature, rather than an active policy tool. Indeed, in the U.S. nation-wide deposit insurance has been in place since the Banking Act in 1933. It can be motivated in an environment with inefficient bank runs (Diamond and Dybvig, 1983), which are not the focus of this paper.
In the following I describe the individual elements of the model in more detail, solve the problem of firms, households, and banks, define the equilibrium and describe how I calibrate the model.

2.1. Firms

Firms operate a production technology and produce output using capital $K_t$ as the only input, 

$$F(K_t) = K_t^\alpha,$$

where $\alpha < 1$. Firms start with an initial capital stock $K_0$. In subsequent periods, capital depreciates with a rate $\delta$ and firms buy new capital from households and banks, such that the capital stock evolves according to

$$K_t = (1 - \delta)K_{t-1} + I_{m, t-1} + \tilde{I}_{b, t-1}.$$ 

Here, $I_{m, t-1}$ denotes newly produced capital by households investing in the financial market, and $\tilde{I}_{b, t-1}$ new capital created through bank loans. Denoting by $R_t$ the return on newly produced capital, firm profits in period $t$ can be written as

$$\pi^f_t = F(K_t) - R_t(K_t - (1 - \delta)K_{t-1}).$$

The firm problem is to maximize expected profits, discounted by the household’s stochastic discount factor $\beta_t$,

$$\max_{K_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \left( \prod_{\tau=0}^{t-1} \beta_\tau \right) \pi^f_t.$$  

(1)
2.2. Household Problem and Liquidity Demand

An infinitely-lived, representative household maximizes her lifetime utility over consumption $C_t$ and liquidity services from deposits $D_t$. Households have a preference for different varieties of bank deposits indexed by $i \in [0, 1]$. Different varieties could represent a bank specializing in online banking, a big international bank with a prestigious brand, or a local bank with personal relations between clients and advisors. Following Dixit and Stiglitz (1977) I model this preference by expressing $D_t$ as a CES composite of varieties $D_{t,i}$,

$$D_t = \left[ \int_0^1 D_{t,i}^{\eta-1} di \right]^{\frac{1}{\eta}}.$$

In this model of monopolistic competition product differentiation gives banks some market power, the degree of which is governed by the elasticity of substitution $\eta$. Higher values of $\eta$ indicate greater ease of substitutability between varieties, implying lower market power. I assume that $\eta > 1$, such that deposits of different banks are substitutes.

Next to deposits, households can invest in the financial market $I_t^m$, to produce capital goods that are sold to firms in the following period. Households are also the owners of firms and banks. Firms rebate their profits $\pi_t^f$ and banks make a net dividend payment $d_t^b$, which may take negative values when raising new equity.

The household’s discount factor $\beta_t$ evolves according to a two-state Markov process. At the beginning of each period, households learn whether $\beta_t = \beta_H$, resulting in high interest rates (state $s = H$), or $\beta_t = \beta_L > \beta_H$, resulting in low interest rates (state $s = L$). The probability of transitioning from state $s$ to $s'$ is denoted $P_{ss'}$.

Utility is linear in consumption $C_t$ and concave in deposits $D_t$, and the problem of the

---

14 Arguably, bank market power is not only driven by product differentiation, and for example customer "stickiness" is likely another important determinant. The advantage of the Dixit-Stiglitz model of monopolistic competition is that it is quite tractable in general equilibrium. It is commonly used in the macro literature, and has recently gained popularity in the banking literature (e.g. Drechsler et al., 2016).
representative household is given by

$$\max_{C_t, I_t^m, D_{t,i}} \mathbb{E}_0 \sum_{t=0}^{\infty} \left( \prod_{\tau=0}^{t-1} \beta_\tau \right) [C_t + \gamma v(D_t)]$$

with

$$D_t = \left[ \int_0^1 D_{t,i}^{\frac{\eta-1}{\eta}} di \right]^{\frac{\eta}{\eta-1}},$$

s.t.

$$C_t + I_t^m + \int_0^1 D_{t,i} di = R_t I_{t-1}^m + \int_0^1 r_{t,i} D_{t-1,i} di + d_i + \pi_t^f - T_t,$$

$$C_t, I_t^m, D_{t,i} \geq 0.$$  \hspace{1cm} (2)

Here, $\gamma \geq 0$ measures the household’s preference for liquidity services, and $v(D_t)$ is a CRRA utility function with relative risk aversion $\theta$. The deposit rate offered by bank $i$ is denoted $r_{t,i}$, and $T_t$ are taxes raised by the government to fund the deposit insurance. The first constraint is the household’s budget constraint, and the second a non-negativity constraint for consumption, deposits and investments in the financial market.

The first-order condition with respect to $I_t^m$ yields the household’s Euler equation

$$R_{t+1} \beta_t = 1.$$  \hspace{1cm} (3)

Since $\beta_t$ can only take two values, this condition implies that the economy is either in a high-rate environment with $R_{t+1} = 1/\beta_H \equiv R_H$, or a low-rate environment with $R_{t+1} = 1/\beta_L \equiv R_L$. This property highlights the analytical attractiveness of the chosen utility function, namely that the return on capital is a function of the current state only.\(^{15}\)

Next to the financial market, households invest in bank deposits. The demand for deposits of bank $i$ is given by the first-order condition with respect to $D_{t,i}$:

$$D_{t,i}(r_{t+1,i}) = \left[ \frac{1 - r_{t+1,i}/R_{t+1}}{\gamma v'(D_t)} \right]^{-\eta} D_t,$$  \hspace{1cm} (4)

where I use that $\beta_t = 1/R_{t+1}$ by (3). Banks can attract more funding, the higher the deposit rate $r_{t+1,i}$ they offer, i.e. the lower the interest margin $R_{t+1}/r_{t+1,i}$. The elasticity of substitution $\eta$ governs how elastic demand is with respect to deposit rates, as greater substitutability makes it easier for households to switch to competitors. Finally, the demand for deposits increases in the preference for liquidity services $\gamma$.

\(^{15}\)Arguably, variations in discount factors are not the main driver behind movements in interest rates. The goal here is not to explain why interest rates are low, and this simple formulation allows to study the implications of low interest rates while preserving tractability.
2.3. The Bank’s Problem

In each period $t$, bank $i$ sets its gross deposit rate $r_{t+1,i}$, and decides how much equity to contribute per unit of deposit, denoted $e_{t,i}$. Setting deposit rates, banks are subject to a zero lower bound constraint that requires $r_{t+1,i} \geq 1$. Moreover, capital regulation requires that $e_{t,i} \geq \bar{e}_t$.

Each bank has access to a single project that I refer to as bank loans. Since $e_{t,i}$ is expressed per unit of deposit, the total investment scale of the project is $I^b_t = (1 + e_{t,i})D_{t,i}(r_{t+1,i})$. With probability $q(m_t)$, the project succeeds and produces one unit of capital per unit of investment, that is sold to firms in the following period. The success probabilities across banks are i.i.d., such that there is no aggregate risk.

In case of failure, the project produces nothing and the bank fails. Depositors of failing institutions are repaid by the deposit insurance fund, but failing banks are not bailed out and cannot continue operating. To keep the total number of banks constant, I assume that each failing bank can be replaced by a new entrant, but that the total number of bank licenses is fixed as mass 1.\footnote{Note that it is indeed optimal for new banks to enter, since banks have market power and earn monopolistic rents.}

The project’s success probability increases in the monitoring intensity $m_t \geq 0$ chosen by the bank. In principle, $q(m_t)$ can be any function with $q'(m_t) \geq 0$, that is bounded by $\lim_{m_t \to \infty} q(m_t) \leq 1$, and $\lim_{m_t \to 0} q(m_t) \geq 0$. For concreteness I use as a functional form the CDF of the standard Gaussian distribution, $q(m_t) = \Phi(m_t)$.\footnote{The advantage is that this function is well behaved and bounded between 0 and 1.} Banks incur a cost $c(m_t) = \psi_1 + \frac{\psi_2}{2}m_t^2$ per unit of investment, which consists of two components. The parameter $\psi_1$ governs the overall cost of operating a bank, such as maintaining a branch network and costs of complying with regulation. The second term depends on the bank’s monitoring intensity $m_t$, and creates a trade-off between risk and return.

Crucially, the bank’s monitoring intensity is not observable and can therefore not directly constrained by financial regulation. Instead, the choice of $m_t$ must be incentive compatible, implying that banks generally do not choose the socially optimal level of risk taking.
A crucial element in the analysis will be the value of the bank’s franchise $V_t$, which generally takes strictly positive values due to the market power of banks.\footnote{When deposit rates are constrained by the ZLB, it may potentially be that the bank’s franchise value turns negative. I do not study this case and focus on equilibria with $V_t \geq 0$.} To define the franchise value, it is useful to write the bank’s problem recursively as

$$V_t = \max_{m_{t,i}, e_{t,i}, r_{t+1,i}} \pi^b_{t,t+1} D_{t,i}(r_{t+1,i}) + q(m_{t,i}) \beta_t \mathbb{E}_t V_{t+1},$$

with

$$\pi^b_{t,t+1} = (q(m_{t,i}) \beta_t [R_{t+1}(1 + e_{t,i}) - r_{t+1,i}] - [e_{t,i} + (1 + e_{t,i}) c(m_{t,i})]),$$

s.t.

$$D_{t,i}(r_{t+1,i}) = \left[ 1 - \frac{r_{t+1,i}}{R_{t+1}} \right]^{\frac{\eta}{\gamma'}} D_t,$$

$$e_{t,i} \geq \bar{e}_t,$$

$$r_{t+1,i} \geq 1.$$ \hfill (5)

Here, $\pi^b_{t,t+1}$ are discounted expected profits per unit of deposits raised at period $t$.\footnote{Hence, $V_t$ is defined as the value of a bank after paying out its entire earnings as dividends. Net dividends transferred to households at time $t$ are $d_t^b = [R_t(1 + e_{t-1,i}) - r_{t,i}] D_{t-1}(r_{t,i}) - [e_{t,i} + (1 + e_{t,i}) c(m_{t,i})] D_t(r_{t+1,i})$, with negative values indicating net equity raised. Note that only the net payout matters, since there are no costs to adjusting dividends (in contrast to, e.g. Begenau (2016)).}

The FOC w.r.t. $e_{t,i}$ shows that banks always choose the minimum amount of equity consistent with the capital requirement, $e_{t,i} = \bar{e}_t$:

$$\frac{\partial V_t}{\partial e_{t,i}} = q(m_t) \beta_t R_{t+1} - (1 + c(m_t)) \leq 0.$$ \hfill (6)

The inequality follows from using that $\beta_t R_{t+1} = 1$, by the household’s Euler equation (3). Banks choose the minimum level of capital because bank equity is privately and socially costly. Households can always invest in the financial market, producing one unit of capital per unit of investment. In contrast, a unit of bank loans has the cost $(1 + c(m_t))$. Hence, investments in the bank’s technology are dominated, making it costly to put equity into banks.\footnote{Note, however, that this does not imply that households are unwilling to hold bank stock. If bank stocks were traded, they would in fact do so at strictly positive values, reflecting the monopolistic rents banks earn. The subtle difference here is between raising new equity and the value of outstanding equity. While outstanding stocks are valuable, bank management would never voluntarily raise new equity funding.}

The first-order conditions with respect to the deposit rate and monitoring intensity jointly
determine \( r_{t+1} \) and \( m_t \):

\[
r_{t+1} = \max \left\{ R_{t+1} \left[ 1 - \frac{\eta}{\eta - 1} \frac{(1 - q(m_t))e_t + (1 + e_t)c(m_t)}{q(m_t)} \right], 1 \right\},
\]

\[ c'(m_t)(1 + e_t)D_t = q'(m_t)\beta_t \left( [(1 + e_t)R_{t+1} - r_{t+1}] D_t + \mathbb{E}_t V_{t+1} \right). \tag{8} \]

In the first case in (7), banks set the deposit rate at an interior solution, proportional to the return on capital. Deposit rates are below \( R_{t+1} \), as banks pass on the cost of equity and charge a mark-up that depends on the elasticity of substitution between deposits \( \eta \) (a higher level of \( \eta \) implies less market power and hence higher deposit rates). If this interior solution is smaller than 1, the ZLB binds and the second case in the max-function applies.

Condition (8) is the bank’s incentive compatibility constraint governing its risk taking. It equates the marginal cost of monitoring on the left-hand side to the marginal benefit on the right-hand side. The higher the bank’s expected continuation value \( \mathbb{E}_t V_{t+1} \) on the right hand side, the more intensely it monitors.

Note that I dropped all \( i \) subscripts, since the solution to the bank’s problem does not depend on any individual characteristics and hence the equilibrium is symmetric.

### 2.4. Government

To close the model, the government runs a balanced budget to finance the deposit insurance. In a symmetric equilibrium, each period a fraction \( (1 - q(m_{t-1})) \) of banks fail, such that the government needs to raise taxes of

\[
T_t = (1 - q(m_{t-1}))r_tD_{t-1}
\]

(9) to repay depositors of failing banks. Finally, the government sets the capital requirement \( \bar{e}_t \), which is taken as exogenously given for now (section 5 derives the welfare-maximizing level of \( \bar{e}_t \)).

### 2.5. Equilibrium

The only state variables of the model are the capital stock \( K_t \) and the realization of the discount factor \( \beta_t \). Both are known at the beginning of the period, and decisions are made subsequently. In the following equilibrium definition I use the fact that the equilibrium is symmetric (since none of the bank’s first-order conditions are bank-specific).
Definition. Given government policies \( \{T_t, \bar{e}_t\}_{t=0}^{\infty} \), transition probabilities \( P_{ss'} \), an initial state \( s_0 \in \{H, L\} \), and an initial capital stock \( K_0 \), a symmetric competitive equilibrium is a set of prices \( \{R_t, r_t\}_{t=0}^{\infty} \) and allocations \( \{K_{t+1}, I_t^m, I_t^b, C_t, D_t, e_t, m_t\}_{t=0}^{\infty} \), such that

(a) Given an initial capital stock \( K_0 \) and prices \( \{R_t\}_{t=0}^{\infty} \), firms maximize profits (1).

(b) Given prices \( \{R_t, r_t\}_{t=0}^{\infty} \) and government policies \( \{T_t, \bar{e}_t\}_{t=0}^{\infty} \), households maximize life-time utility solving (2).

(c) Given prices \( \{R_t\}_{t=0}^{\infty} \) and government policies \( \{T_t, \bar{e}_t\}_{t=0}^{\infty} \), banks maximize net dividends solving (5).

(d) Market clearing is satisfied at any time \( t \geq 0 \)

- aggregate resource constraint:

\[
C_t + I_t^m + I_t^b(1 + c(m_t)) = F(K_t),
\]

- capital:

\[
K_t = (1 - \delta)K_{t-1} + I_t^m + q(m_t)I_{t-1}^b,
\]

with

\[
I_t^b = (1 + e_t)D_t.
\]

The set of equations describing the equilibrium is summarized in appendix A.1. The forward-looking nature of the bank’s problem and the occasionally binding ZLB constraint potentially complicate solving the model. However, owing to the simple stochastic structure and linear utility function, the equilibrium values of all variables relevant for the bank’s forward-looking problem \( (R_t, e_t \text{ and } D_t) \) depend on the current state only, i.e. they are memory-less and independent of the time period \( t \).\(^{21}\) This property simplifies solving the bank’s problem, since it allows to explicitly derive the expected franchise value as

\[
E_s V_{t+1} = P_{ss'} V_s + (1 - P_{ss'}) V_{s'}.
\]

For ease of notation I denote the value of a memory-less variable \( x_t \) in state \( s \) simply as \( x_s \), and the expectations given state \( s \) as \( E_s x_t+1 = E_t[x_{t+1}|s] \).

\(^{21}\)In fact, the equilibrium values of all variables except for \( I_t^m \) and \( C_t \) are memory-less.
2.6. Discussion of the Framework

To preserve tractability there are several shortcuts and assumptions in the model, on which I elaborate here.

The money in the utility approach assumes a social value of bank debt. While a shortcut, the banking literature has identified several micro-foundations that motivate this assumption. Because bank debt is information-insensitive, it protects depositors from better informed traders (Gorton and Pennacchi, 1990; Dang et al., 2017). Its demandability may incentivize monitoring (Diamond, 1984), and facilitates the transformation of risky long-term assets into liquid and safe claims (Diamond and Dybvig, 1983; Ahnert and Perotti, 2017). Moreover, banks invest into an ATM network and electronic payment infrastructure that make deposits a convenient medium of exchange.

I assume that the banks’ technology to produce new capital goods is less efficient than the financial market. This assumption ensures that it is socially costly to provide equity to banks, as the financial market provides a superior outside option, i.e. that the banks’ FOC w.r.t. equity (6) is negative. It ensures a meaningful trade-off between overcoming moral hazard and the cost of bank equity.

One might think that deposit insurance and the liquidity service of deposits already make bank equity costly. However, this is only true in a model in which the balance sheet size of a bank is fixed, such that the relevant opportunity cost is the interest paid on deposits. If instead banks can expand the size of their balance sheet, the relevant opportunity cost is the required return of shareholders, which in this model is given by the financial market.22

It is less important that banks do not have access to the dominating “financial market” technology. In the Internet Appendix I derive an extension of the model, in which the market technology dominates yet banks choose to invest in their own “bank loan” technology. This can be ensured by controlling the overall cost ψ1 of operating a bank. Here I simply assume that banks only have access to their own technology, allowing for a streamlined solution of the model.23

---

22To see this, consider a version of the model in which the total investment size \( I^b_t = (1 + \epsilon)D_t \) is fixed at \( I^b_t = 1 \), s.t. \( D_t = \frac{1}{1+\epsilon_t} \). In this case, the banks’ FOC w.r.t. \( \epsilon_t \) is negative if and only if \( r_{t+1} < \frac{R_{t+1}}{q(m_t)} \). Given risk-neutrality, the required return on deposits is indeed less than \( R_{t+1}/q(m_t) \) if there is either deposit insurance or deposits have a liquidity service in the utility.

23In the extension, the market technology succeeds with probability \( \mu < 1 \), and one needs to check that in
Since households are risk-neutral, it does not matter that the financial market represents a riskless investment. What matters is that there is some risk in the bank’s investment, to introduce the risk shifting moral hazard problem.

In the model banks are never bailed out upon failure, and bank failures impose no distress costs on the rest of the system. While in the real world regulators tend to take a tough no-bailout stance ex ante, the immense costs of the failure of large institutions often forces regulators to bail out banks ex post (e.g. Acharya and Yorulmazer, 2007). Both elements could easily be added to model. A (possibly endogenous) bailout probability would dampen the franchise value effect, as banks may be able to continue operating even after failure. This would increase the risk taking incentives of banks. A higher cost of bank failure would motivate tighter capital regulation in equilibrium.

In the model, the total number of bank licenses is limited to 1, ensuring that entry is always profitable since there are rents to be earned. Indeed, in the real world there are many regulations that limit entry to the banking sector. For example, the process of obtaining a bank charter is relatively long and complicated, and requires approval from the Office of the Comptroller of the Currency the FDIC, and in some cases the Federal Reserve.²⁴

Nevertheless, one may expect some consolidation after a prolonged period of near-zero interest rates. Appendix B.2 presents evidence on the evolution of concentration in the banking industry consistent with this notion. Since the ZLB started binding in 2008, the average number of banks per county has been dropping from around 14.5 to 13.5.

This effect is not present in my model, because entry is still profitable at the ZLB, as banks earn monopolistic profits.²⁵ Allowing for some consolidation may lighten the negative effect of the zero lower bound on franchise values. However, the problem at the ZLB is not that banks compete too fiercely with each other, but that banks face competition from cash as an alternative source of liquidity. At the ZLB, deposit rates are fixed at the corner solution \( r_t = 1 \), independently of the degree of competition between banks. Therefore, it is not clear

---

²⁴See for example the FAQ on “How can I start a bank?” on the website of the Board of Governors of the Federal Reserve System.

²⁵Except when rates are so low that \( V_t < 0 \). At this point banks scale down until the return on capital recovers sufficiently. Weak investment then pushes output below potential, and the economy ends up in a classic liquidity trap. I study this case in a previous version of the paper, available upon request.
that lower concentration would substantially support bank profitability.

2.7. Calibration

I derive as many results as possible analytically, but also rely on a numerical solution of the model when analytics are ambiguous. For this purpose, I calibrate the model to U.S. data. The calibration also allows me to get a sense for the magnitude of the effects I find, and allows me to quantify optimal capital requirements in section 5.

I think of the high-rate state as “normal times”, with safe, short term rates away from the ZLB, such as the period from the 1990s until the financial crisis in 2008. Accordingly, I set $\beta_H = 0.95$ to generate a return on capital of around 5.5%, which is equal to the average yield on AAA corporate bonds over the period 1996-2008, as reported in FRED. The level of rates in the low-rate state is one of the main comparative statics of interest. In the baseline calibration, I set $\beta_L = 0.975$ to target the average AAA corporate bond yield over 2009-2013 at around 2.5%.

Regarding aggregate Macro moments, I set $\delta = 0.065$, equal to the average depreciation rate of the U.S. capital stock from 1970-2016, computed using the BEA’s Fixed Assets Tables 1.1 and 1.3. Using the same data and period, I compute an average capital-output ratio of 3.25. Accordingly, I set $\alpha = 0.38$, such that $K_H/Y_H = 3.25$ in the high-rate state.

Next, I set the capital requirement s.t. $\bar{e}_t/(1 + \bar{e}_t) = 0.085$, equal to the minimum requirement for the Tier 1 capital ratio in the Basel III framework.

To calibrate the cost function parameters I first set $\psi_1 = 0.018$. This is equal to the median bank’s ratio of non-interest expenses to assets, calculated using Call reports data over the period 1984-2013. The parameter $\psi_2$ is the cost of monitoring. Hence, it governs bank risk taking, and I set it to get an equilibrium $q(m_H)$ around 0.9925 in the high-rate state. This way, the success probability is in line with the average annual proportion of banks failing in the U.S., computed to be 0.76% by Davydiuk (2017) using the fail bank list issued by the FDIC.

The elasticity of substitution $\eta$ affects bank market power and hence interest margin $R_H - r_H$. Following Drechsler et al. (2016) I use call reports data to proxy deposit rates as the

\[ BofA \text{ Merrill Lynch US Corporate AAA Effective Yield [BAML0A1C0AAEY], retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/BAML0A1C0AAEY.} \]
deposit interest expense per unit of deposits. Similarly, I calculate the interest income rate as the ratio of interest income over total assets, and the interest margin as the difference between interest income and expense ratio. The average interest margin over the period 1996-2008 is 3.5%, consistent with a value of $\eta = 5$.

Given the calibration of the bank variables I set the parameters $\gamma$ and $\theta$ governing the preference for liquidity. Doing so, I target the ratio of aggregate deposit liabilities of U.S. chartered institutions to the aggregate debt instruments of non-financial corporates using data from the Flow of Funds. This is only one moment to set two parameters. Accordingly, I first restrict $\theta = 1$, the log case, and then set $\gamma = 0.003$ to get a ratio ratio of $D_H/(D_H + K_H^m) = 0.2$, consistent with the Flow of Funds data.

Finally, in the baseline calibration I set the transition probabilities equal to $P_{HH} = 0.9$ and $P_{LL} = 0.8$. This implies an expected duration of 10 years spent in the high-rate state, and 5 years in the low-rate state. For comparison, the Federal Funds Rate target range was at 0% for seven years, from December 2008 - December 2015. All parameter values are summarized in table 1 in appendix A.2.

3. First Best and Inefficiencies

Before studying the equilibrium properties of the model, it is useful to characterize the first best allocation (FB), and contrast it to the competitive equilibrium (CE). The first best allocation is the solution to a planner’s problem, who directly chooses risk taking, consumption and investment subject to aggregate resource constraints:

$$\max_{C_t, I_t^m, D_{t,i}, m_{t,i}, e_t} \mathbb{E}_0 \sum_{t=0}^{\infty} \left( \prod_{\tau=0}^{t-1} \beta_\tau \right) [C_t + \gamma v(D_t)]$$

with

$$D_t = \left[ \int_0^1 \frac{\eta_{t,i}^{-1}}{\eta_{t,i}} di \right]^{\eta_{t,i}},$$

s.t.

$$C_t + I_t^m + \int_0^1 (1 + e_{t,i}) D_{t,i} (1 + c(m_{t,i})) = F(K_t),$$

$$K_t = (1 - \delta) K_{t-1} + I_{t-1}^m + q(m_t)(1 + e_t) D_t$$

$$C_t, I_t^m, m_{t,i}, D_{t,i} \geq 0$$

From the CES aggregator it follows immediately that the planner allocates the same amount of deposit funding to each bank, $D_{t,i} = D_t$. Moreover, since the bank’s investment technology
is dominated, investments into banks can only be socially useful if they are in the form of deposits, and hence \( e_t = 0 \). The remaining variables are chosen according to the first-order conditions w.r.t. \( I^m_t, m_t \) and \( D_t \):

\[
\beta_t R_{t+1} = 1 \quad (11)
\]

\[
c'(m_t) = q'(m_t) \quad (12)
\]

\[
D_t = \left( \frac{\gamma}{1 - q(m_t) + c(m_t)} \right)^\theta \quad (13)
\]

These three conditions are readily compared to their counterparts in the competitive equilibrium. First, (11) is equivalent to the household’s Euler equation (2), implying that the overall level of capital accumulation is not distorted.

In contrast, condition (12) differs from its counterparts in the CE. In the FB allocation, \( \frac{c'(m_t)}{q'(m_t)} = 1 \). This is not generally true in the CE, as revealed by the bank’s FOC w.r.t monitoring (8).

Similarly, condition (13) can be compared to the demand for deposits by households (13), after rewriting it as

\[
D_t = \left( \frac{\gamma}{1 - r_{t+1}/R_{t+1}} \right)^\theta . \quad (14)
\]

Clearly, the quantity of deposits in the CE is only equal to its FB level if \( \frac{r_{t+1}}{R_{t+1}} = c(m_t) - q(m_t) \). However, this is not generally true, see (7).

These two comparisons show that misallocations arise because banks do not choose the optimal amount of risk taking, and do not provide the optimal amount of liquidity via deposits. These inefficiencies are a result of the frictions in the model that are (i) deposit insurance and limited liability, and (ii) monopolistic competition.

Limited liability and deposit insurance give banks an option-like payoff as they do not internalize the losses incurred in case of failure. This convex payoff structure induces excessive risk taking. On the other hand, (ii) monopolistic competition implies that banks may take less risk relative to the FB. The reason is that the bank’s franchise value reflects rents due to market power. These rents are of private value to banks but do not add to welfare. Generally, banks trade off the gains from shifting risk on the deposit insurance against the risk of loss of franchise value. In the baseline calibration, banks take excessive risk relative to the first best.
4. Risk Taking off and at the Zero Lower Bound

Given that banks do not take the optimal level of risk, the goal of this section is to study how risk taking incentives are affected by the level of interest rates and other parameters in the model. It turns out that the answer crucially depends on whether the ZLB binds. For that reason, I start by showing under what conditions banks indeed do become constrained in setting their deposit rate.

4.1. Zero Lower Bound

Banks set their deposit rate according to the first-order condition (7). This may either be at an interior solution if the return of capital is sufficiently high, and at the corner solution \( r_{t+1} = 1 \) if the ZLB binds.

Lemma 1. At any time \( t \), in state \( s \) the ZLB is slack (i.e. banks set an interior deposit rate \( r_s \geq 1 \)) if and only if

\[
\beta_s \leq \beta_s^{ZLB},
\]

where \( \beta_s^{ZLB} \) is implicitly defined by

\[
\beta_s^{ZLB} = 1 - \frac{\eta}{\eta - 1} \left( 1 - q(m_s^*) \right) \bar{e_s} + (1 + \bar{e_s}) c(m_s^*),
\]

and \( m_s^* \) denotes the equilibrium level of monitoring as a function of \( \beta_s^{ZLB} \), implicitly defined by (3), (4), (7) and (8).

Lemma 1 defines a threshold \( \beta_s^{ZLB} \), below which the ZLB binds. In the baseline calibration, this threshold is around 0.965 in both the high-rate and low-rate state, such that deposit rates hit the ZLB when the return on capital drops below 3.5% (\( = 1/\beta_s^{ZLB} - 1 \)). The return on capital is above this threshold in the high-rate state (5.5%), while in the low-rate state \( R_L = 2.5\% \) and the ZLB binds.

4.2. Do Low Interest Rates Spur Risk Taking?

The following proposition answers this question, by examining the comparative statics of equilibrium monitoring with respect to a marginal change in the discount factor \( \beta_t \).

Proposition 1. Hold \( \beta_{t+1}, \beta_{t+2}, \ldots \) fixed. The comparative statics of monitoring \( m_t \) with respect to the discount factor \( \beta_t \) depend on whether the ZLB binds:
• If $\beta_t \leq \beta_s^{ZLB}$ (slack ZLB) a marginal increase in $\beta_t$ increases equilibrium monitoring:
\[ \frac{dm_t}{d\beta_t} \geq 0 \]
(i.e. lower rates induce less risk taking).

• If $\beta_t > \beta_s^{ZLB}$ (binding ZLB), a marginal increase in $\beta_t$ (falling rates) unambiguously decreases equilibrium monitoring $m_t$:
\[ \frac{dm_t}{d\beta_t} \leq 0 \]
(i.e. lower rates induce more risk taking).

To see this result, rewrite the bank’s FOC w.r.t. monitoring (8) as
\[ \frac{d}{dt} \left( m_t \right) + \frac{1}{(1 + \epsilon_t)} \frac{r_{t+1}}{R_{t+1}} + \beta_t \frac{E_t V_{t+1}}{(1 + \epsilon_t) D_t} \]
\[ \frac{m_t}{m_t} \frac{R_{t+1}}{R_{t+1}} \]
\[ \text{Margin on current loans} \quad \text{Discounted franchise value} \]
The left-hand side increases in monitoring $m_t$. The right-hand side reveals that the level of interest rates affects monitoring via a margin channel and a discounting channel. A higher level of $\beta_t$ implies that banks discount their continuation value $E_t V_{t+1}$ less, boosting overall franchise value. Via this discounting channel, lower interest rates induce banks to monitor more intensely, i.e. take less risk.

On the other hand, an increase in $\beta_t$ directly pushes down $R_{t+1}$ (by the households Euler equation 3). A low return on capital may harm interest margins and thereby induce higher risk taking. However, as long as banks set deposit rates according to the interior solution in (7), the relevant ratio $\frac{r_{t+1}}{R_{t+1}}$ on the right-hand side is not a function of $\beta_t$:
\[ \frac{r_{t+1}}{R_{t+1}} = 1 - \frac{\eta (1 - q(m_t)) \bar{e}_t + (1 + \bar{e}_t) c(m_t)}{q(m_t)} \].
\[ \text{Intuitively, market power allows banks to pass on a reduction in } R_{t+1} \text{ to depositors, such that the relative interest margin } \frac{r_{t+1}}{R_{t+1}} \text{ remains stable. In contrast, when the ZLB binds,} \]
\[ \frac{r_{t+1}}{R_{t+1}} = \frac{1}{R_{t+1}} \]
and any reduction in $R_{t+1}$ directly eats into interest margins. This effect is illustrated in the left panel of figure 4, which plots the return on capital $R_L$ and equilibrium deposit rate $r_L$

27To be precise, the left-hand side increases in $m_t$, as long as $m_t \geq \frac{1}{2}$. In equilibrium, always $m_t \geq \frac{1}{2}$ because the Gaussian distribution is symmetric, and .

23
in the low-rate state, against the level of the discount factor $\beta_L$. As long as $\beta_L \leq \beta_L^{ZLB}$, banks can decrease deposit rates proportionately, guaranteeing a stable interest margin. In contrast, when $\beta_L > \beta_L^{ZLB}$ the ZLB binds and margins shrink.

The overall effect on risk taking depends on the balance between the discounting and margin channel. The right panel of figure 4 plots the equilibrium success probability $q(m_L)$ against the discount factor $\beta_L$. The discounting effect dominates as long as the ZLB is slack ($\beta_L \leq \beta_L^{ZLB}$), as margins are stable. The effect on risk taking is rather modest; success probabilities rise by a few basis points as the return on capital falls from above 5% (at $\beta = 0.95$) to around 3.5% (at $\beta = \beta_L^{ZLB}$).

In contrast, the margin effect dominates when the ZLB binds. This can be shown analytically (proposition 1), by plugging $r_{t+1} = 1$ into (16) (also using 14). Figure 4 suggests that the magnitude of the effect is sizable, more than doubling the annual probability of failure, from around 0.7% to above 1.5%, as the return on capital falls from 3.5% (at $\beta_L = \beta_L^{ZLB}$) to 2% (at $\beta_L = 0.98$).

---

28Note that the comparative statics in proposition 1 regard a marginal change in $\beta$, keeping $\beta_{t+1}, \beta_{t+2}, \ldots$ constant. Figure 4 complements this result, varying the level $\beta_L$ that applies whenever the low-rate state obtains.

29Analytically, this can be shown by plugging the interior solution from (7) into (16), also using (14). Doing so, the right hand side of (7) only depends on $\beta$ via the discounting effect.
Figure 4 also reveals that a binding ZLB in the low-rate state affects risk taking in the high-rate state ($s = H$, see dashed red line). Even though in the high-rate state the ZLB is slack, the possibility of a binding ZLB in the future erodes franchise value. Risk taking incentives are not only affected by current profits, but also expected profitability going forward.

4.3. Expectations Matter

Because expected future profitability affects franchise value, it matters for how long the economy is expected to remain at the ZLB:

**Proposition 2.** Suppose that $\beta_{H} < \beta_{ZLB}^{H}$ and $\beta_{L} > \beta_{ZLB}^{L}$ (ZLB slack in the high-rate state, and binding in the low-rate state). There exists a threshold $\hat{\beta} \leq \beta_{ZLB}^{L}$, s.t. if

$$\beta_{L} \geq \hat{\beta},$$

then $V_{H} > V_{L}$. In this case, equilibrium monitoring in states $s = H, L$ decreases the more time the economy is expected to spend at the ZLB:

$$\frac{d m_{s}}{d P_{LL}} < 0, \quad \frac{d m_{s}}{d P_{HH}} > 0$$

When $\beta_{L} > \hat{\beta}$, the ZLB binds and intermediation margins are sufficiently compressed, such that $V_{H} > V_{L}$. In this case, the overall value of banks is lower, the more time the economy spends in the low-rate state. Low expected profitability erodes franchise value and boosts risk taking incentives.

The left panel in figure 5 illustrates the result of proposition 2. It plots the equilibrium success probability $q(m_{s})$ against the likelihood of remaining in the low-rate state $P_{LL}$. In the baseline calibration indeed $V_{H} > V_{L}$, such that condition (19) is satisfied and an increase in $P_{LL}$ results in more risk taking (lower $q(m_{s})$).

The right panel in figure 5 connects this result to the yield curve, here calculated assuming that the expectations hypothesis holds.\(^{30}\) This calculation shows that a zero interest environment may be particularly problematic if the yield curve flattens substantially and rates are expected to be at the ZLB for long. The target range for the Fed Funds rate was lowered to 0% in December 2008, where it remained for seven years, until the Fed started lifting rates in December 2015. An expected duration of seven years corresponds to a probability of staying

\(^{30}\)I.e. the forward rate from date $t$ to $t + \tau$ is calculated as $R_{t,t+\tau} = (R_{t+1} \times R_{t+2} \cdots \times R_{t+\tau})^{1/\tau}$.\[^{25}\]
in the low-rates state of around $P_{LL} = 0.85$. In the Eurozone rates are expected to remain near-zero for an even longer time. The ECB lowered its deposit facility rate close to zero by the beginning of 2009, and did not start the process of increasing rates by end 2017.

Even with rates in the U.S. rising, the overall level of interest rates is expected to remain low (perhaps due to demographic change and weak demand for finance by corporations). This increases the likelihood that upon the next major macroeconomic shock rates will hit the ZLB again. Proposition 2 also shows that even when banks are not currently constrained by the ZLB, the prospect of a binding ZLB in the future affects incentives. The more likely the economy transitions from the high-rate to the low-rate state (lower $P_{HH}$), the greater the chance that banks face weak profitability in the future, and hence the more risk they take.

### 4.4. Discussion of the Mechanism

In the model, risk taking is driven by bank franchise value, consistent with several empirical studies. For example, Jiang et al. (2017) exploit the differential process of bank deregulation across U.S. states to show that a deregulation-induced increase in competition increases risk taking through reduced profits and bank franchise values. Similarly, Beck et al. (2013) find support for a positive relation between bank competition and fragility across a large set of countries.
Franchise value, in turn, is driven by interest margins and bank competition. As long as the ZLB is slack, interest margins are determined by market power. At the ZLB, bank competition is distorted, as depositors are unwilling to accept negative interest rates.\(^{31}\)

This approach allows me to focus on the distinct effect of the ZLB on bank competition and franchise values. It is consistent with high-frequency studies of bank stock price reactions to monetary policy announcements. English et al. (2012) and Ampudia and Van den Heuvel (2017) find that interest rate decreases boost bank stock prices when the ZLB is slack. Ampudia and Van den Heuvel (2017) also show that the effect reverses during the recent period with near-zero interest rates.

The overall mechanism mirrors the evidence in Heider et al. (2016). In a diff-in-diff setting, the authors show that negative policy rates in the Eurozone have eaten relatively more into the interest margin of banks with more deposit relative to wholesale funding. Consistent with the notion that tight margins spur risk taking, these banks are shown to lend to riskier borrowers.

5. Capital Regulation and Other Policy Options

In the model, the main policy tool to curb moral hazard is capital regulation. Given that the zero lower bound on deposit rates induces banks to take more risk, one might expect that the optimal reaction to near-zero interest rates is to tighten capital requirements. I find exactly the opposite. After calculating optimal capital requirements, I show that this finding can be explained by an effect I label “capital overhang”.

5.1. Optimal Capital Regulation

I calculate the welfare-maximizing, state-dependent levels of the capital requirement \(\{c_H^*, c_L^*\}\), that maximize the representative household’s expected lifetime utility. For this purpose, I

\(^{31}\)Drechsler et al. (2016) argue more generally that the closer interest rates are to zero, the more bank deposits compete with cash. If I were to introduce a more general substitutability between cash and deposits, a reduction in interest rates would undermine bank market power even further away from zero. Consequently, the margin channel described in proposition 1 would already be at play with a slack ZLB, and a reduction in interest rates might increase risk taking incentives even when the ZLB is slack. Still, incentives would be affected disproportionately once the ZLB binds.
simulate the model for 300,000 random paths of length of 200 years, starting in the high-rate state \((s_0 = H)\). I then pick the combination of capital requirements that maximizes the average lifetime utility across the 300,000 draws, and define the resulting allocation as the “second best”. To be very clear, this means that I take deposit insurance and the level of competition as given, i.e. they are not part of the policy choice set.\(^{32}\)

The result of this exercise is presented in the top panel of figure 6, for different levels of \(\beta_L\). When \(\beta_L < \beta^{ZLB}_L\) interest rates are high and the ZLB is slack. In this region, I find an optimal capital requirement between 7% and 8% in both the low-rate and high-rate state, somewhat above the level currently required according to the Basel III regulatory framework.\(^{33}\) In contrast, when the ZLB binds \((\beta_L > \beta^{ZLB}_L)\) the optimal capital requirement in the low-rate state, \(e^*_L\), is U-shaped in \(\beta_L\), while that in the high rate state, \(e^*_H\) increases.

To understand this pattern, note that in equilibrium banks take too much risk, i.e. the equilibrium level of monitoring \(m_L\) is below the first best (bottom-left panel of figure 6). A higher capital requirement curbs the incentives to take excessive risk:

**Proposition 3.** An increase in the capital requirement induces banks to monitor more intensely in equilibrium:

\[
\frac{dm_s}{de_s} \geq 0.
\]

The intuition for this result is the typical “skin in the game” mechanism. As shareholders put more of their own funds at stake, their payoff becomes less convex, inducing a more prudent investment strategy (Holmstrom and Tirole, 1997).

While an increase in the capital requirement reduces risk taking, it also undermines the

\(^{32}\)Nation-wide deposit insurance is a long established institution that has been in place in the U.S. since the Banking Act in 1933.

\(^{33}\)In the model, the only assets of banks are risky loans and the requirement is expressed as a fraction of total non risk-weighted assets. Strictly speaking, the capital requirement therefore resembles more closely the leverage ratio requirement of Basel III, rather than capital requirements, which are risk-weighted. At the same time, in the model banks only invest in risky loans, which tend to carry relatively high regulatory risk-weights. The quantitative assessment of the optimal capital requirement can therefore be interpreted as a leverage requirement on risky loans, somewhere between the leverage and capital requirements of Basel III. According to Basel III capital regulation, banks are required to hold Tier 1 plus Additional Tier 1 capital of 6%, plus an additional 2.5% in the “Capital Conversation Buffer”, all as a fraction of risk-weighted assets (BIS, 2011). Moreover, Basel III requires a “leverage ratio” of at least 3% of Tier 1 capital over total (non risk-weighted) assets.
Figure 6: The top panel plots the optimal state-dependent capital requirement, for different levels of the discount factor in the low-rate state. The bottom panels plot for both the first best and second best the equilibrium success probabilities $q_L$ and quantity of deposits $D_L$. The vertical dotted line marks the threshold $\beta_L^{ZLB}$, beyond which the ZLB binds in the low-rate state. Parameters are calibrated as described in section 2.7.

liquidity creation of banks. Focusing for now on the case of a slack ZLB ($\beta_L \leq \beta_L^{ZLB}$), the bottom-right panel of figure 6 reveals that banks provide too little liquidity, i.e. the equilibrium level of $D_L$ is below first best.

From equation (18), banks pass on the cost of capital in the form of a higher interest margin (for a given level of $m_L$, $r_L/R_L$ decreases in $e_L$). The cost of capital in (18) consists of two components. First, a higher capital requirement reduces the implicit subsidy from deposit insurance $(1 - q(m_s))e_s$. Second, the operating cost per unit of investment $(1 + e_s)c(m_s)$. As banks pass on both cost components, the demand for liquidity by households declines as $e_L$ increases.

The optimal level of the capital requirement trades off a reduction in risk taking against
lower liquidity provision. These two countervailing forces are optimally traded off at a capital requirement near 8% when $\beta_L = \beta_H = 0.95$ and the ZLB is slack. Recall from proposition 1 that lower discount rates induce banks to take less risk. This explains that for higher levels of $\beta_L$ the optimal capital requirement decreases slightly, allowing for a higher level of liquidity provision while keeping the equilibrium success probability at a stable level.

This trade-off changes substantially when deposit rates are constrained by the ZLB ($\beta_L > \beta^{ZLB}_L$). Fixing $r_L = 1$ implies that banks can no longer pass on the cost of capital to households, such that deposits become more attractive for households. As a result, the gap to the first best level of liquidity provision $D_L$ closes at the ZLB. In fact, for very low levels of interest rates deposits become so attractive that households consume even more liquidity than in the first best.

At the ZLB, the optimal capital requirement in the low-rate state, $e^*_L$ drops substantially, while $e^*_H$ increases. As evident in the bottom-left panel of figure 6, the drop in the optimal capital requirement occurs despite higher risk taking. As I show next, this pattern is explained by an “capital overhang” effect.

### 5.2. Capital Overhang

Proposition 3 shows that tighter capital requirements weaken risk shifting incentives. A novel result here is that at the ZLB this “skin in the game” effect may be overruled by a countervailing effect, that I refer to as “capital overhang” effect. When banks are unable to pass on the cost of capital to depositors, tight capital requirements eat into bank profitability. This is shown in the following lemma:

**Lemma 2 (Capital Overhang).** For a given level of monitoring $m_s$, bank profits as a function of the capital requirement are given by

$$\pi^b_s(e_s; m_s) = \begin{cases} \frac{1}{1-\eta} [(1 - q(m_s))e_s + (1 + e_s)c(m_s)], & \text{if } \beta_s \leq \beta^{ZLB}_s \\ q(m_s)(1 - \beta_s) - c(m_s) - e_s[1 + c(m_s) - q(m_s)], & \text{if } \beta_s > \beta^{ZLB}_s \end{cases}$$

Clearly,

- $\frac{\partial \pi^b_s(e_s; m_s)}{\partial e_s} \geq 0$ if $\beta_s \leq \beta^{ZLB}_s$ (ZLB slack)
- and $\frac{\partial \pi^b_s(e_s; m_s)}{\partial e_s} \leq 0$ otherwise.


Since weak profitability implies lower franchise value, capital regulation becomes a less effective tool to curb risk shifting incentives at the ZLB. This is shown in equilibrium in the numerical solution in figure 7, using the baseline calibration of section 2.7. The left panel plots the equilibrium franchise value in the low-rate state $V_L$ against the capital requirement $\bar{e}_L$ (keeping $\bar{e}_H$ fixed), for different levels of $\beta_L$ and likelihood of remaining in the low-rate state. At $\beta_L = 0.95$ the ZLB is slack, and capital requirements have an overall positive effect on $V_L$, in line with lemma 1. In contrast, when the ZLB binds (at $\beta_L = 0.975$), higher capital requirements erode profitability. The figure also reveals that the adverse effect on franchise value is particularly strong, the longer the economy remains at the ZLB in expectation. Intuitively, the longer the ZLB binds, the lower is bank profitability in expectation.

The right panel of figure 7 shows the implication for risk taking. The more the capital requirement depresses franchise values, the less it curbs risk shifting. For example, franchise values drop much more on the dashed black line with circle markers (representing $P_{LL} = 0.95$, or an expected duration of 20 years at the ZLB) than on the green, dashed line (representing $P_{LL} = 0.8$, or an expected duration of 5 years at the ZLB). Accordingly, in the right panel the black line is flatter, i.e. a marginal increase in capital requirements reduces risk shifting incentives relatively less.

In fact, in the limiting case with $P_{LL} = 1$, the capital overhang effect completely overrules the skin in the game effect, such that capital regulation does not affect risk taking at all. This
result can be shown analytically:\footnote{To see that the capital requirement becomes ineffective, evaluate (16) at $s = L$ and $r_L = 1$, using $D_L$ from (14) and $V_L$ from (20). After some algebra, it can be seen that with $P_{LL} = 1$ all $\bar{e}_L$ drop out from the right hand side of (16), implying that $m_L$ is unaffected by $\bar{e}_L$.}

**Proposition 4.** Suppose $\beta_L > \beta_L^{ZLB}$ (ZLB binds at the low-rate state). In the limiting case $P_{LL} = 1$ (the ZLB binds forever), equilibrium monitoring $m_L$ is unaffected by the level of capital requirements,

$$\frac{dm_L}{d\bar{e}_L} = 0.$$ 

### 5.3. Implications for Macro-Prudential Policy

The capital overhang effect renders capital requirements less effective in curbing moral hazard, motivating a weaker use at the ZLB. This effect explains the drop in the optimal capital requirement $e_L^*$ for $\beta_L > \beta_L^{ZLB}$ found in figure 6.\footnote{The optimal level of the capital requirement drops as low as zero at the ZLB. A 100\% leverage prescription is certainly an over-statement of the quantitative magnitude of the effect, and reveals a weakness of the model. The binary shock structure ignores the “loss absorbing capacity” element of bank capital. Under a more realistic distribution of returns the bank’s failure probability would rise exponentially as its equity tends to zero, resulting in a strictly positive optimal level of the capital requirement. Nevertheless, the capital overhang effect would still be present in this alternative version of the model. The binary shock structure puts the focus on moral hazard and still delivers a realistic quantitative assessment of the optimal capital requirement away from the ZLB.} It also explains why capital regulation optimally allows more risk taking at the ZLB (bottom left panel of figure 6).

At the same time, the marginal return to monitoring is higher at lower levels of $m_t$, i.e. $q(m_t) - c(m_t)$ is concave. This opposing effect motivates a tighter capital requirement, explaining the overall U-shaped pattern of $e_L^*$. When intermediation margins are still relatively high, the capital overhang effect dominates. At some point, risk taking is so strong that the marginal return to monitoring is very high and it becomes optimal to again increase the capital requirement. However, the optimal capital requirement at $\beta_L = 0.975$ is still substantially lower than at $\beta_L = \beta_L^{ZLB}$ ($\approx 4.5\%$ vs $7\%$).

Figure 6 further reveals that depressed franchise values in the low-rate state motivate a tighter level of capital requirement $e_H^*$. Because the ZLB is slack in the high-rate state, here the effectiveness of capital requirements is not undermined by the capital overhang effect.
Heightened risk taking incentives induced by low profitability in the low-rate state therefore unambiguously motivate a higher level of $e^*_H$.

The model thus offers a novel rationale for counter-cyclical capital regulation. To the extent that interest rates are low in bad times, counter-cyclical capital regulation may be motivated because at the ZLB it is a less effective tool to curb moral hazard.

Recent contributions show that a regulatory leverage limits may be motivated in models with welfare-relevant pecuniary externalities (e.g. Lorenzoni, 2008; Stein, 2012; Korinek and Simsek, 2016). In the policy debate, a common rationale is that buffers built up in good times should be available to be used in bad times (e.g. Goodhart et al., 2008). In contrast, the argument here is based purely on the effectiveness of capital requirements to overcome risk shifting problems. Thereby, it even applies in a setup without any costs to raising equity, and absent welfare-relevant pecuniary externalities.

Another implication of the “capital overhang” effect is that monetary- and macro-prudential policy cannot be seen in isolation. In the policy debate it is sometimes argued that monetary policy is too blunt a tool to target financial stability (e.g. Bernanke, 2015). However, this argument sees monetary policy as an independent, alternative tool to macro-prudential regulation. If instead near-zero interest rates undermine the effectiveness of prudential policies, there may be a case to jointly set monetary and prudential policy.

5.4. An Alternative Policy

Is there a better policy response than merely adjusting capital requirements at the ZLB? I consider as an alternative policy a subsidy $\tau_t$ per unit of deposits, paid to banks whenever the ZLB binds. The subsidy is set to replicate whatever negative rate banks want to set, effectively eliminating the ZLB. That is,

$$\tau_t = \min \{1 - r_{t+1}, 0\}.$$ 

To finance the subsidy, the government raises lump sum taxes of $\tau_tD_t$.

The subsidy resembles cheap funding schemes such as the ECB’s Targeted Long Term Refinancing Operations (TLTRO), or the Bank of England’s Funding for Lending scheme. These policy schemes provide cheap funding to banks, conditional on making loans. For example, the second round of TLTRO (TLTRO-II) was announced in March 2016, and enabled
banks to borrow up to 30% of the amount of their existing stock of loans to non-financial corporations and households. The maximum rate banks have to pay for TLTRO-II funding is zero (ECB, 2017).

The subsidy effectively eliminates the ZLB constraint for banks. Accordingly, it restores bank profitability and hence incentives, as illustrated in the left panel of figure 8. This figure highlights the difference between the competitive equilibrium with and without the subsidy, and a counter-factual economy absent the ZLB friction. The left panel shows that under the subsidy the risk taking of banks is much lower than without it, close to the level of risk taking in an economy without the ZLB friction.

However, the overall effect of the subsidy on welfare is ambiguous. The right panel plots a welfare gap, defined as the percentage deviation of the representative household’s lifetime utility from the first best.\(^{36}\) Only when rates are quite low ($\beta_L$ high) does the subsidy result in a higher level of welfare. The taxes raised to fund the subsidy create an additional distortion. From the view of depositors, investments in deposits are too attractive, such that the overall demand for liquidity is inefficiently high.

\(^{36}\)As in the previous section, I simulate the model for 300,000 random paths of length of 200 years, and calculate the average lifetime utility across all draws.
6. Conclusion

Since the 1980s real interest rates across advanced economies have followed a steady downward trend. While many central banks are preparing for a “normalization” of interest rates, low rates are likely here to stay (Summers, 2014). This new environment of near-zero interest rates requires re-thinking some fundamental questions across macro- and financial economics. This paper is a step in this direction, highlighting potential consequences for banking and financial regulation.

I show that the zero lower bound distorts bank competition, weakens profitability and induces banks to take more risk in a search for yield. This effect is particularly strong the yield curve flattens substantially and rates are expected to remain low for a long time. And even after monetary policy “normalization”, incentives are affected if the ZLB is expected to bind again in the future.

While the ZLB has often been discussed as a constraint to monetary policy, I show that it also affects the effectiveness of prudential regulation. At the ZLB, banks cannot pass on the cost of capital to depositors, such that tight capital requirements erode franchise value. This “capital overhang” has the perverse effect of increasing risk taking incentives, undermining the usual “skin in the game” effect of capital regulation (Holmstrom and Tirole, 1997).

An implication is that in an environment of structurally low interest rates, monetary and macro-prudential policies cannot be seen in isolation. Hence, future research should study their joint determination. As I work in a real model, I also leave unanswered how a struggling banking system interacts with inflation expectations. And while I do not model entry, another relevant question is how industry structure changes if rates remain low for a long time.
References


Bech, M. L. and A. Malkhozov (2016). How have central banks implemented negative policy rates?


Bernanke, B. (2015). Should monetary policy take into account risks to financial stability?


Drechsler, I., A. Savov, and P. Schnabl (2017a). Banking on deposits: Maturity transformation without interest rate risk.


A. Paper Appendix

A.1. Equilibrium conditions

All equilibrium conditions can be summarized as follows:

- **Firms**

  \[
  F(K_t) = K_t^\alpha, \\
  K_t = (1 - \delta)K_{t-1} + I_t^m + q(m_t)I_{t-1}^b, \\
  I_t^b = (1 + e_t)D_t, \\
  \alpha K_t^{(\alpha-1)} = R_t - (1 - \delta).
  \]

- **Households**

  \[
  R_{t+1}\beta_t = 1, \\
  D_t = \left(\frac{\gamma}{1 - r_{t+1}/R_{t+1}}\right)^\theta, \\
  C_t = F(K_t) - I_t^m - I_t^b(1 + c(m_t)).
  \]

- **Banks**

  \[
  \frac{c'(m_t)}{q'(m_t)} = \frac{[(1 + e_t)R_{t+1} - r_{t+1}] D_t + \bar{E}_t V_{t+1}}{(1 + e_t)D_t R_{t+1}}, \\
  V_t = \pi_{t,t+1}^b + q(m_{t,i})\beta_t \bar{E}_t V_{t+1}, \\
  \pi_{t,t+1}^b = (q(m_t)\beta_t [R_{t+1}(1 + e_t) - r_{t+1}] - [e_t + (1 + e_t)c(m_t))] D_t, \\
  e_t = \bar{e}_t, \\
  r_{t+1} = \max \left\{ R_{t+1} \left[ 1 - \frac{\eta}{\eta - 1} \frac{(1 - q(m_t))e_t + (1 + e_t)c(m_t)}{q(m_t)} \right], 1 \right\}.
  \]
A.2. Calibration

The following table summarizes the calibration of the model and data sources.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Target Moment</th>
<th>Data Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta_H = 0.95$</td>
<td>Average corporate bond yield 1996-2008, $R_H = 1.055$</td>
<td>FRED</td>
</tr>
<tr>
<td>$\beta_L = 0.975$</td>
<td>Average corporate bond yield 2009-2013, $R_L = 1.025$</td>
<td>FRED</td>
</tr>
<tr>
<td>$\delta = 0.065$</td>
<td>Average depreciation rate of U.S. capital stock 1970-2016</td>
<td>BEA Fixed Assets Tables</td>
</tr>
<tr>
<td>$\alpha = 0.38$</td>
<td>Average U.S. capital-output ratio 1970-2016, $K_H/Y_H = 3.25$</td>
<td>BEA Fixed Assets Tables</td>
</tr>
<tr>
<td>$\bar{e}_s = 0.0929$</td>
<td>Basel III bank capital requirement, $\bar{e}_s/(1 + \bar{e}_s) = 0.085$</td>
<td>BIS</td>
</tr>
<tr>
<td>$\psi_1 = 0.018$</td>
<td>Median U.S. bank’s non-interest expense / assets 1984-2013, $\psi_1 = 0.018$</td>
<td>Call Reports (obtained through WRDS)</td>
</tr>
<tr>
<td>$\psi_2 = 0.0018$</td>
<td>Average annual failure rate of U.S. banks, $q(m_H) \approx 0.9924$</td>
<td>Davydiuk (2017)</td>
</tr>
<tr>
<td>$\eta = 5$</td>
<td>Average interest margin of U.S. banks from 1996-2013 $R_H - r_H = 3.5%$</td>
<td>Call Reports (obtained through WRDS)</td>
</tr>
<tr>
<td>$\gamma = 0.003$</td>
<td>Deposit liabilities of U.S. chartered institutions / debt instruments of non-financial corporates, $D_H/(D_H + I_H^{m}) = 0.2$</td>
<td>Flow of Funds</td>
</tr>
<tr>
<td>$P_H = 0.9$</td>
<td>Expected duration in high-rate state of 10 years</td>
<td>N/A</td>
</tr>
<tr>
<td>$P_L = 0.8$</td>
<td>Expected duration in low-rate state of 5 years</td>
<td>N/A</td>
</tr>
</tbody>
</table>

A.3. Proof of Proposition 2

This appendix shows (i) that $V_H > V_L$ when $\beta_L < \hat{\beta}$, and (ii) that in this case equilibrium monitoring increases in $P_{HH}$ and decreases in $P_{LL}$.

(i) Use the definition of $V_t$ and $\pi_{t,t+1}$ from (5), and that $\mathbb{E}_s V_{t+1} = P_{ss'} V_s + P_{ss'} V_{s'}$, to find the franchise value of the bank in state $s \in \{H, L\}$:

$$V_s = \frac{1}{\Lambda} \left[ (1 - q(m_{s'})\beta_{s'} P_{s's'}) \pi_s D(r_s) + q(m_s) \beta_s P_{ss'} \pi_{s'} D(r_{s'}) \right]$$, \hspace{1cm} (20)

with

$$\Lambda \equiv (1 - q(m_H)\beta_H P_{HH})(1 - q(m_L)\beta_L P_{LL}) - (q(m_H)\beta_H P_{HL})(q(m_L)\beta_L P_{LH})$$,
and $D(r_s)$ is defined in (14). By lemma 1, if $\beta_L > \beta^ZLB_L$, the ZLB binds. In this case, one can write $\pi_L$ as

$$\pi_L = q(m_L) \left[ (1 + \bar{e}_L) - \frac{1}{R_L} \right] - [\bar{e}_L + (1 + \bar{e}_L)c(m_L)].$$

Moreover, with a binding ZLB at $r_L = 1$,

$$D(1) = \left( \frac{\gamma}{1 - 1/R_L} \right)^{\theta}.$$

Clearly, $\lim_{R_L \to 1} \pi_L < 0$, and $\lim_{R_L \to 1} D(1) = \infty$. Hence,

$$\lim_{R_L \to 1} \pi_L D(1) = -\infty.$$

Inspecting (20), it is clear that the term $\pi_L D(r_L)$ has a greater weight on $V_L$ than $V_H$ (since $(1 - q(m_H)\beta_H P_{HH}) > q(m_H)\beta_H P_{HL}$). Hence, $V_L$ tends faster to $-\infty$ as $\beta_L$ increases and there is a threshold $\hat{\beta} < 1$ s.t. for $\beta_L > \hat{\beta}$ it must be that $V_H > V_L$.

(ii) From (16), monitoring increases in $E_t V_{t+1}$. With $V_H > V_L$ it follow immediately that $E_s V_{t+1} = P_{ss} V_s + P_{ss'} V_{s'}$ increases in $P_{HH}$ in the high state, and decreases in $P_{LL}$ in the low state. Hence, monitoring increases in $P_{HH}$ and decreases in $P_{LL}$.
**B. Additional Evidence**

**B.1. Interest Margins and Deposit Rates at the ZLB**

Figure 1 from the introduction shows that the spread between safe corporate bonds and the deposit expense ratio has declined since 2009. The left panel of figure 9 complements this data by showing the spread between interest income and deposit interest expense ratio of the median U.S. bank in the call reports data. Analogously to the interest expense ratio, the income ratio is defined as total interest income (riad4107) divided by total assets (rcfd2170).

As in figure 1, a compression in spreads is visible in these series too, though the magnitude of the drop is smaller and occurs slightly earlier - perhaps because non-performing loans started pushing down bank interest income already in 2007.

That interest income ratios are somewhat more stable than the return on safe bonds in figure 1 is consistent with the notion that banks start lending to riskier borrowers (since riskier borrowers pay higher interest rates). It is also driven by the fact that bank assets have relatively long maturity, so that margins only come under pressure once their long-term assets roll off. Drechsler et al. (2017a) show that banks in the U.S. lengthened the duration of their balance sheets during the zero-lower-bound period, which has limited the compression of their net interest margins.

In my model I cannot study these gradual effects as loans are re-priced every period. Nevertheless, the comparison to highly rated corporate bonds in figure 1 shows that for a given level of risk margins on new business are significantly compressed since 2009.

The right panel of figure 9 shows for a longer horizon the spread between the rate on 30 year mortgages (as reported in FRED), and the median deposit interest expense ratio. I calculate the mean of this spread for three phases: 1985 - 1995, 1996 - 2007, and 2007 - 2013.

In 1994 the Riegle-Neal Interstate Banking and Branching Efficiency Act removed several obstacles to banks opening branches in other states and provided a uniform set of rules regarding banking in each state. This act increased competition, with an evident negative effect on interest margins. In 2008 the ZLB starts binding, explaining the second drop in margins, analogous to the left panel and figure 1.

This pattern of interest margins is consistent with the model. Away from the ZLB, margins are determined by the level of competition (parameter $\eta$ in the model). When the ZLB binds,
the market power of banks breaks as depositors face cash as an attractive outside option. Accordingly, a further compression in margins occurs.

**Deposit Rates** Figure 10 expands on figure 2 in the introduction. This more comprehensive perspective shows that the skewness and concentration of the distribution is a phenomenon particular to the ZLB period after 2009. This is despite substantial swings in the Federal Funds rate over the relevant period.

**B.2. Evolution of Bank Concentration**

A central prediction of the model is that the ZLB distorts bank competition, as cash provides an attractive alternative source of liquidity for households. In the light of weakening profitability, one may expect the industry to consolidate.

Figure 11 presents evidence of the evolution of bank concentration since 1994, using branch-level data on deposit holdings from the FDIC. The left panel shows that the aggregate number of banks has been steadily decreasing since 1994. In contrast, the average number of banks per county increases from around 13 in 1994 to almost 14.5 in 2008. These trends are consistent with the interpretation that after 1994 competition between banks increased. In 1994 the Riegle-Neal Interstate Banking and Branching Efficiency Act removed several obstacles to interstate-banking. This allowed the most efficient banks to venture into other states, explaining the increase in the average number of banks per county. At the same time, less efficient
Figure 10: For the years 1994-2013, this figure plots the cross-sectional distribution of deposit interest expense ratios across U.S. banks in the call reports data. The deposit interest expense ratio is defined as interest expenses per unit of deposits.
banks leave the market, explaining the decrease in the number of banks on the national level.

As the ZLB starts binding in 2008, banks again face fiercer competition. However, this
time tighter competition is not the result of fiercer competition with each other, but a result
of the fact that depositors have cash as an alternative source of liquidity with zero net return.
Accordingly, the growth in the number of banks per county reverses, falling in tandem with
the aggregate number of banks, and almost all the way back to its 1994 level. Likely other
drivers behind the fall in the number of banks are the emergence of online banking and fintech,
as well as bank failures triggered by the financial crisis.

The right panel of figure 11 further supports this interpretation by plotting deposit Herfindahls on a national and the country level. Following Drechsler et al. (2016), I calculate the
county-level Herfindahl by summing the deposit holdings across all branches of a bank in a
given county, and then calculating the Herfindahl as the sum of squared deposit market shares
of all banks in a county. Analogously, I calculate the aggregate Herfindahl by summing the
deposit holdings across all branches of a bank in the entire U.S.

Unsurprisingly, the Herfindahls have an inverse relationship to the number of banks, con-
firming that county-level concentration decreases from 1994-2008, but then starts increasing
again as the ZLB binds from 2009 onwards. Interestingly, by 2015 the mean County Herfindahl
surpasses its 1994 level.

Figure 11: The left panel plots the number of banks on a nation-wide level (left axis), and the mean number
of banks per county (right axis). Analogously, the right panel plots Herfindahl based on bank-level
deposits on a nation-wide level, and per county.
B.3. Non-Interest Income and Expense

Even if banks are unable to set negative interest rates on deposits, they may be able to do so effectively by increasing fees. While this is certainly a possibility, non-interest income of banks is small relative to their interest income. The median U.S. bank has an average net interest income of 3.9% over the period 1984-2013. Over the same period, average non-interest income is only 0.6%.

Looking separately at service charges, fees amount to an even smaller number of around 0.37% per unit of deposits. As shown in figure 12, this number has actually been coming down in recent years, to below 0.25%. Hence, both magnitude and trend suggest that fees do not make up for squeezed interest margins.