Alternative Models of Interest Rate Pass-Through in Normal and Negative Territory

Mauricio Ulate
Federal Reserve Bank of San Francisco

September 2020

Working Paper 2020-31

https://www.frbsf.org/economic-research/publications/working-papers/2020/31/

Suggested citation:

The views in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Federal Reserve Bank of San Francisco or the Board of Governors of the Federal Reserve System.
Alternative Models of Interest Rate Pass-Through in Normal and Negative Territory

Mauricio Ulate*

September 9, 2020

Abstract

In the aftermath of the Great Recession, many countries used low or negative policy rates to stimulate the economy. These policies gave rise to a rapidly growing literature that seeks to understand and quantify their impact. A fundamental step when studying the effectiveness of low and negative policy rates is to understand their transmission to loan and deposit rates. This paper proposes two models of pass-through from policy rates to loan and deposit rates that can match important stylized facts while remaining parsimonious. These models can be used to study the transition between positive and negative policy rates and to quantify the impact of negative rates on banks.


Keywords: Negative Interest Rates, ZLB, Monetary Policy, Bank Profitability.

*Email: mauricio.ulate@sf.frb.org. The views in this paper do not necessarily reflect the views of the Federal Reserve Bank of San Francisco or the Federal Reserve System. For useful comments I thank Jose Vasquez, Ashley Lamquist, Rupal Kamdar, Yuriy Gorodnichenko, Andrés Rodríguez-Clare, Walker Ray, Juan Herreño, and participants in various seminars and conferences. All errors are my own.
1 Introduction

During the Great Recession of 2008-2009 many countries cut their policy rates to zero (or its vicinity) to fight the downturn and stimulate the economy. The slow recovery that followed the recession featured nominal rates that remained at zero in many advanced countries and even became negative in others. The effectiveness of these low and negative rates has been debated in the press, central banks, and the academic literature, but the matter remains unsettled. A fundamental issue when studying low or negative policy rates is their transmission to other interest rates that play an important role in the broader economy. Two such rates are the interest rate that commercial banks charge on loans (hereafter referred to as the “loan rate”), and the interest rate that commercial banks pay their customers for deposits (hereafter referred to as the “deposit rate”). The pass-through of the policy rate to loan and deposit rates is a crucial component in determining the effectiveness of cutting the policy rate in low or negative territory.

Empirically, papers like Drechsler et al. (2017) have found that the pass-through of the policy rate to deposit rates is between 0.5 and 0.6 when rates are in their normal range, while papers like Eisenschmidt and Smets (2019) have documented that this pass-through is close to zero when rates are very low or negative. For loan rates, Altavilla et al. (2019) and Ulate (2019) have documented a pass-through of between 0.5 and one when rates are in their normal range. The value of the loan rate pass-through when rates are low or negative is a more contested issue, with papers like Eggertsson et al. (2019) and Amzallag et al. (2019) claiming that the pass-through is close to zero (or negative), and papers like Ulate (2019) and Eisenschmidt and Smets (2019) finding that it is still positive. Even though there are disagreements in this literature, and the topic is still evolving, a rough consensus of the facts is that the pass-through of the policy rate to loan and deposit rates is positive but incomplete (say between 0.5 and 0.8) in normal times, the pass-through of the policy rate to the deposit rate is roughly zero in negative territory, and the pass-through of the policy rate to the loan rate is intermediate in negative territory.

In this paper, I propose two models of interest rate pass-through that can capture the facts mentioned in the previous paragraph while remaining tractable. These models extend and modify the static banking model of Ulate (2019), which is unable to capture non-unitary pass-through. The original model of Ulate (2019) contains separate borrowers and savers that solve a two period problem. Additionally, it assumes that customers: 1) choose a single bank from a continuum of possibilities over which they have differentiated preferences, 2) choose their bank before other quantities of interest, 3) have a unitary intertemporal elasticity of substitution, and 4) can only save/borrow through banks. These four assumptions imply
Figures 1: Behavior of rates in the original model of Ulate (2019)

Notes: This figure shows the loan rate, policy rate, and deposit rate behavior as a function of the policy rate in the original model of Ulate (2019).

that customers have CES preferences between banks in loan demand and deposit supply. As a consequence, banks set the loan rate as a mark-up on the policy rate and the deposit rate as a mark-down on the policy rate during “normal times” (i.e. when the policy rate is above a threshold of roughly 50 basis points). These mark-ups and mark-downs are almost constant, generating a pass-through of the policy rate to the deposit rate and the loan rate which is essentially one. This complete pass-through (illustrated in Figure 1) during normal times is inconsistent with the stylized facts mentioned above.

The first extension developed in this paper maintains assumptions 1, 2 and 4, but deviates from Ulate (2019) by relaxing assumption 3. Specifically, I assume that borrowers and savers have a CES utility function between today and tomorrow with an intertemporal elasticity of substitution greater than one. Consequently, this extension is denoted “High Intertemporal Substitution” model. For borrowers, an intertemporal elasticity of substitution greater than one means that when rates are high they want to borrow a small share of their income. This gives lenders “less monopoly power” and makes them charge a smaller loan spread. In contrast, savers want to save a higher share of their income when rates are high,
which means that deposit-taking banks have “more monopoly power” and charge a higher deposit spread. This leads to a behavior of rates, illustrated in Figure 2, which is consistent with the stylized facts about pass-through discussed earlier.

The second extension developed in this paper maintains assumptions 1, 2 and 3, but deviates from Ulate (2019) by relaxing assumption 4. Specifically, savers are allowed to use three type of instruments: cash, deposits, and bonds. Furthermore, cash and deposits (combined through a CES aggregator) provide liquidity, which is valued by customers. Consequently, this extension is denoted “Liquidity and Bonds” model. This setup implies that the choice of how many deposits to maintain is determined by the comparison of the price of deposits to the price of liquidity. In contrast, the choice of bank is determined by comparing a bank’s price of liquidity with the price of liquidity offered by other banks. The total amount of deposits supplied to a bank is a combination of the amount that its customers want to deposit and the probability that a given customer chooses that bank. Hence, banks face two different margins of substitution, one given by the elasticity between deposits and cash in the liquidity aggregator, and the other one given by the dispersion in preferences.
across banks. This leads to a behavior of rates which I illustrate in Figure 3. The behavior of the loan rate is identical to the one in the first model, but the behavior of the deposit rate is different, as this variable grows linearly with the policy rate.\footnote{The “Liquidity and Bonds” model relies on the same mechanism as the “High Intertemporal Substitution” model to obtain a non-unitary pass-through for borrowers, but uses a completely different mechanism to obtain a non-unitary pass-through for savers, as explained in the text. That is why the behavior of the loan rate is the same in both models but the behavior of the deposit rate is different.}

After discussing the models and their implications for loan and deposit rate pass-through, I proceed to discuss their implications for the Return on Equity (ROE) of banks. The behavior of ROE under the three models is displayed in Figure 4. While the pattern of ROE is not exactly the same in the extended models as in the original model of Ulate (2019), the behavior is not too different. In all three models, ROE falls steeply with the policy rate below a certain threshold $\tilde{\iota}$, but has a more moderate behavior above the threshold.\footnote{The value of the threshold and the reason for its existence are discussed extensively in Ulate (2019) and will also be covered in Section 2. A second threshold $\check{\iota}$ is also present but will not be discussed as much in this paper, since it is likely to be below -1.5%.}
Figure 4: Model implied relationship between ROE and $i$

Notes: This figure describes the model-implied relationship between bank (gross) return on equity ($F'/F$, denoted ROE), on the y axis, and the policy rate ($i$), on the x axis for three different models. The levels $\tilde{i}$ and $\hat{i}$ represent thresholds where commercial banks start reacting differently to the policy rate; their expressions are given in Section 2. The model of Ulate (2019), explained in Section 2, is the black line. The “High Intertemporal Substitution” extended model of Section 3 is the blue line. Finally, the “Liquidity and Bonds” extended model of Section 4 is the red line.

In the static model of Ulate (2019) the slope of ROE below the threshold is around 5, while above the threshold it is around 1. In the “High Intertemporal Substitution” model, the slope above the threshold is around zero at first and eventually becomes greater than one. In the “Liquidity and Bonds” model, the slope above the threshold is slightly higher than in the original model throughout. Even though there are slight differences, the overall behavior of ROE is similar under models that feature non-unitary pass-through. This serves to reassure the reader that the results in Ulate (2019) are not reliant on the assumption of (approximately) unitary pass-through.

The extended models proposed in this paper can be used to study the impact of negative nominal interest rates on banks. Since they feature a more realistic pass-through in “normal times”, they can also be used to build models that more seamlessly capture the transition between positive and negative territory. Even though both extended models produce a similar pass-through in normal times, there is still value in having two alternative models, since researchers might want to include or exclude alternative saving vehicles (i.e., cash or bonds) in their DSGE models depending on their specific purposes.
To my knowledge, this is the first paper that develops banking models with non-unitary pass-through that contain a continuum of banks and monopoly power. Papers like Drechsler et al. (2017), Wang et al. (2019), Kurlat (2019), or Balloch and Koby (2019) have developed models where the pass-through of the policy rate to the deposit rate in normal territory can differ from one. Certain parametrizations of those models can produce a deposit pass-through in the 0.5 to 0.6 range. These models are related to the “Liquidity and Bonds” model, as agents can save not only via deposits with banks, but also in cash or bonds. Cash and deposits provide liquidity services, while bonds do not. When rates are low, and bonds and money have a similar return, deposits are not very useful and banks have little monopoly power, so they set small spreads.

In the papers mentioned in the previous paragraph, the mechanism for non-unitary pass-through relies on having a limited number of banks. These banks, which have significant size, realize that they affect the aggregate deposit rate, which changes their rate setting behavior. If these same models are modified to have a medium or large number of banks (or, in the limit, a continuum), then even with the introduction of bonds and cash, the pass-through of the policy rate to the deposit rate approaches one. The timing assumption in the “Liquidity and Bonds” model, where customers must chose their bank before their saving amount, is what separates the second extended model from previous papers.

To illustrate the importance of the point in the previous paragraph, consider the model of Drechsler et al. (2017). For a given set of parameter values, and a single bank ($N = 1$), the pass-through of a cut in the policy rate from 2% to 1% is exactly half (i.e., 0.5). However, if all parameters are kept fixed but the number of banks is increased to five ($N = 5$), the same measure of pass-through increases from 0.5 to 0.93. This means that even with a medium, yet realistic, number of banks, the model approximately delivers unitary pass-through. For a researcher that is interested in a local labor market with a small number of banks, these types of models can be useful to capture non-unitary pass-through. In contrast, for a researcher trying to calibrate a DSGE model at the national level, introducing a realistic number of banks would lead to a nearly complete pass-through in this family of models. Another downside of having a finite number of banks, is that assumptions must be made about the evolution of the number of banks in order to be able to solve the model. In this paper, I develop models that feature non-unitary pass-through during normal times while featuring a continuum of banks, so that the setup remains tractable and can be used to analyze a national economy in a general equilibrium setup.

3Specifically, the parameter values used here are $\delta = 1$, $\eta = 1.1$, $\varepsilon = 2$ and $\rho = 0.5$. Other parameter choices would deliver different values for the pass-through, but the overall message that increasing $N$ pushes the pass-through towards unity would remain intact.
This paper is related to the theoretical literature that studies the usefulness of negative or low policy rates, like Brunnermeier and Koby (2018), Eggertsson et al. (2019), Rognlie (2016), Balloch and Koby (2019), Sims and Wu (2019), de Groot and Haas (2020), and Wang (2019), while being more limited in scope. While the current paper does not present any empirical results, it is motivated by the empirical literature that discusses the effectiveness of low and negative nominal interest rates. This literature includes papers like Ampudia and van den Heuvel (2019), Borio et al. (2017), Claessens et al. (2018), Basten and Mariathasan (2018), Eisenschmidt and Smets (2019), Lopez et al. (2020), Heider et al. (2019), Hong and Kandrac (2018), etc. The paper is also related to the literature that studies the pass-through of policy rates to other rates in normal times, including papers like Drechsler et al. (2017), Drechsler et al. (2018), Wang et al. (2019), Kurlat (2019).

The rest of the paper proceeds as follows. Section 2 briefly describes the static model of Ulate (2019) and its implications for interest rate pass-through and bank ROE. Section 3 describes the “High Intertemporal Substitution” model, its assumptions, setup, and intuition. Similarly, Section 4 describes the “Liquidity ans Bonds” model. Finally, Section 5 concludes.

2 The Static Model of Ulate (2019)

In the model of Ulate (2019), there is a continuum of banks, indexed by \( j \), between zero and one. Each bank is granted an amount of equity \( F_j \) as an endowment at the beginning of the period, which it combines with an amount of deposits \( D_j \). On the asset side, banks issue loans \( L_j \) and hold reserves \( H_j \). Banks seek to maximize their resources at the end of the period, once loans and deposits have been repaid. Each bank faces a downward-sloping loan demand and an upward-sloping deposit supply captured through a CES aggregator.

The maximization problem that individual bank \( j \) faces is the following:

\[
\max_{i^d_j, L_j, i^d_j, D_j, H_j} (1 + i^l_j)L_j + (1 + i)H_j - (1 + i^d_j)D_j
\]

s.t.

\[
L_j = \left( \frac{1 + i^l_j}{1 + i^d_j} \right) L
\]

\[
D_j = \begin{cases} 
(\frac{1 + i^l_j}{1 + i^d_j})^{-\varepsilon} D & \text{if } i^d_j \geq 0 \\
0 & \text{if } i^d_j < 0 
\end{cases}
\]

\[
L_j + H_j = F_j + D_j
\]

\[
H_j \geq 0.
\]
Equation (1) represents loan demand. Equation (2) represents deposit supply, and it indicates that a bank obtains no deposits if it sets negative nominal deposit rates. Equations (1) and (2) can be derived directly from the behavior of borrowers and savers using the four assumptions mentioned in the introduction, as illustrated in Appendices A.1-A.3 of Ulate (2019). The aggregate amounts of loans demanded by firms and deposits supplied by households are $L$ and $D$ respectively. These aggregate quantities are assumed to be unaffected by any rates in this partial equilibrium model, but they can be made endogenous in more elaborate general equilibrium models. Equation (3) is the bank balance sheet constraint, indicating that total assets (loans plus reserves) have to equal liabilities (which are just deposits) plus equity. Equation (4) states that reserves at the central bank must be nonnegative.

This model assumes that $\varepsilon^l > 1$ and $\varepsilon^d < -1$, that all banks are given the same amount of initial equity $F_j = F$, and that $D > L > F$. The formal solution to the bank problem is described in Ulate (2019), here I provide a brief summary. The solution consists of regimes that apply depending on the level of the policy rate. Regime 1 applies when $i \geq \tilde{i}$, Regime 2 when $i \leq i < \tilde{i}$, and Regime 3 when $i < \tilde{i}$. The thresholds are given by $\tilde{i} \equiv -\frac{1}{\varepsilon^d} > 0$ and

$$\tilde{i} = \frac{(\frac{L}{F})^{\frac{1}{\varepsilon^l}} - \frac{1}{\varepsilon^l - 1} + 1}{1 + \frac{1}{\varepsilon^l - 1} \frac{L}{F} - \frac{1}{\varepsilon^l} \frac{L}{F} - \frac{1}{\varepsilon^l - 1}} < 0.$$

In Regime 1, when the policy rate is in “normal” territory, all banks set the same gross loan and deposit rates, which are given as a mark-up and mark-down on the gross policy rate:

$$1 + i^l_j = \frac{\varepsilon^l}{\varepsilon^l - 1} (1 + i), \quad 1 + i^d_j = \frac{\varepsilon^d}{\varepsilon^d - 1} (1 + i).$$

This is reminiscent of the solution of the pricing problem of a monopolistically competitive good producer. In this model, the absolute values of $\varepsilon^l$ and $\varepsilon^d$ will be high in order to match the steady state spreads between the loan rate and the policy rate and between the policy rate and the deposit rate. Consequently, the values of $\frac{\varepsilon^l}{\varepsilon^l - 1}$ and $\frac{\varepsilon^d}{\varepsilon^d - 1}$ will be close to one, and pass-through will be nearly complete. As mentioned in Ulate (2019), in this regime all banks obtain an amount of deposits equal to the aggregate deposit supply ($D$), give an amount of loans equal to the aggregate demand of loans ($L$), and hold a positive amount of reserves.

In Regime 2, when $\tilde{i} \leq i < \tilde{i}$, all banks set $i^d_j = 0$, receive an amount of deposits $D$, give an amount of loans $L$, and still hold a positive amount of reserves at the central bank. In this regime the loan rate setting behavior of banks is the same as in Regime 1, since the marginal use of commercial banks’ resources is still as reserves at the central bank. Intuitively, Regime 2 exists because there is a range of low and negative policy rates where
banks prefer to receive deposits even if they earn a low or negative spread on them, because it allows them to maintain their leverage and earn more on their loan franchise. Regime 2 stops applying when the policy rate crosses the threshold \( i < 0 \), where offering deposits at a zero rate is so costly that at least one commercial bank has incentives to stop doing so. Regime 3, which applies when \( i < i \), is described in detail in Ulate (2019), but will not be discussed here.

The behavior of interest rates with respect to the policy rate in this model is described in Figure 1. Since the policy rate is in both axes, the orange line is simply the diagonal. Additionally, it is clear that the loan rate is a mark-up over the policy rate and the deposit rate is a mark-down over the policy rate. Moreover, the spreads are essentially constant when the policy rate is above \( i \) (which is around 50 basis points). The \( x \) axis in Figure 1 only contains realizations of the policy rate that are above \( i \), since \( i \approx -2\% \) in this model. The behavior of return on equity (ROE) is depicted in Figure 4 with a black line. The interest rate \( i \) represents the threshold where further cuts in the policy rate would turn deposit rates negative in the absence of the deposit ZLB. However, since deposit rates are constrained by zero, \( i \) instead represents the point where lowering the policy rate further starts affecting banks disproportionately, because they cannot charge their usual spread on deposits.

The nearly complete pass-through displayed by this model in “normal territory” makes it unable to match the stylized facts described in the introduction. In the following sections I modify this model in order to capture a non-unitary pass-through in normal times.

3 “High Intertemporal Substitution” Model

As mentioned in the introduction, the first extended model relaxes the assumption of a unitary intertemporal elasticity of substitution and instead assumes that agents (both borrowers and savers) have an elasticity of substitution between today and tomorrow which is greater than one. This seemingly small change has profound implications for the loan and deposit pass-through. For borrowers, it means that when rates are high they do not want to borrow much. For savers, in contrast, it means that when rates are high they want to save a lot. Consequently, high rates amplify the monopoly power of banks on the deposit side, but decrease it on the loan side. Therefore, high rates lead to small loan spreads and high deposit spreads, allowing this model to capture the stylized pass-through facts mentioned in the introduction.

Agents in this model choose their bank before their allocations (i.e. consumption and saving/borrowing), and they have a preference shock (with an extreme value distribution) across different banks. Having agents choose their bank before their allocations captures frictions like switching costs or limited attention spans. These frictions correspond to the realistic feature that customers usually choose their bank once and stick with it for long periods of time. Additionally, assuming a preference shock across banks captures the fact
that due to idiosyncratic or geographical characteristics, certain customers might prefer a
given bank for reasons orthogonal to its interest rates.

The assumptions in this model imply that deposit supply and loan demand for a given
bank contain two different elasticities, one related to the elasticity of substitution between
today and tomorrow, and another one related to the elasticity of substitution between dif-
ferent banks (stemming from the preference shock). If the intertemporal elasticity is greater
than the elasticity across banks, increases in the policy rate will increase the loan elastic-
ity but decrease the deposit elasticity, leading to smaller loan spreads but higher deposit
spreads. In the following subsections, I describe the problem of the saver, the problem of
the borrower, and the problem of the bank respectively.

3.1 The Problem of the Savers

A representative saver has CES preferences between today and tomorrow characterized by
the following utility function:

\[ U(C_0, C_1) = \left( \alpha^d \right)^{\frac{1}{\theta^d}} C_0^{\frac{\theta^d-1}{\theta^d}} + \left( 1 - \alpha^d \right)^{\frac{1}{\theta^d}} C_1^{\frac{\theta^d-1}{\theta^d}} \],

where \( C_0 \) is consumption today, \( C_1 \) is consumption tomorrow, \( \alpha^d \) is the importance of con-
sumption today, and \( \theta^d \) is the elasticity of substitution between today and tomorrow. The
saver has income \( \bar{Y}^d \) today, and no income tomorrow. Therefore, he must save in order to
consume tomorrow. Saving can only be done in a continuum of banks between 0 and 1.

An individual bank is indexed by \( j \). Bank \( j \) offers a deposit rate \( i_j^d \). Savers must first
choose the bank that they will put their savings into, and then the amount that they will
save. The budget constraint of the saver, conditional on the choice of bank \( j \), is given by:

\[ C_0 + \frac{C_1}{1 + i_j^d} = \bar{Y}^d. \]

The solution to this problem is:

\[ C_0 = \alpha^d \left( \frac{1}{p_j^d} \right)^{-\theta^d} \frac{\bar{Y}^d}{p_j^d}, \quad C_1 = (1 - \alpha^d) \left( \frac{1}{p_j^d(1 + i_j^d)} \right)^{-\theta^d} \frac{\bar{Y}^d}{p_j^d}, \]

where

\[ p_j^d \equiv \left( \alpha^d + (1 - \alpha^d) \left( \frac{1}{1 + i_j^d} \right)^{1-\theta^d} \right)^{1-\theta^d} \]
is the price index of aggregate consumption for a saver that chooses bank $j$. The indirect utility function of this consumer is $v_d^j = \ln(Y_d) - \ln(p_d^j)$.\footnote{After adding a logarithm to the utility function, which does not alter the maximization problem.}

Up to now, the quantities being discussed are conditional on choosing bank $j$. The next step is to characterize the choice of bank, which is the first stage of the decision process. I assume that the bank choice stage can be described by a stochastic utility approach, where the total utility of choosing a given bank is the sum of the indirect utility obtained in the second stage, and a stochastic component that varies across banks.\footnote{As mentioned earlier, this stochastic component can be due to several things: geographic variation, switching costs, recommendations of family or friends, etc.}

Mathematically:

$$V_d^j = v_d^j + \mu^d \epsilon_d^j,$$

where $\mu^d$ is a positive constant and $\epsilon_d^j$ is a random variable with zero mean and unit variance.

Assuming that the $\epsilon_d^j$ random variable is independently and identically distributed with type-one extreme value distribution, the probability of choosing bank $j$ is given by:

$$Pr_{r_j}^d = Pr(V_d^j = \max_r V_r^d) = \frac{e^{v_d^j/\mu^d}}{\int_0^1 e^{v_d^r/\mu^d} dr} = \frac{(p_d^j)^{1/\epsilon_d^j}}{\int_0^1 (p_r^d)^{1/\epsilon_d^j} dr},$$

as in McFadden (1973). Substituting $1/\mu^d$ for $\epsilon_d^j - 1$, the previous expression becomes:

$$Pr_{r_j}^d = \frac{(p_d^j)^{1-\epsilon_d^j}}{\int_0^1 (p_r^d)^{1-\epsilon_d^j} dr} = \left(\frac{p_d^j}{p^d}\right)^{1-\epsilon_d^j},$$

where $p^d$ is the usual price index: $p^d = \left(\int_0^1 (p_r^d)^{1-\epsilon_d^j} dr\right)^{1/1-\epsilon_d^j}$. This indicates that the probability of choosing a given bank is determined by the ratio of the price of aggregate consumption offered by that bank over the price of aggregate consumption offered by the “average” bank, with an “elasticity” $\epsilon_d^j$ which captures how sensitive the probability is to deviations from the average price. However, what matters for banks is not only the probability that they are chosen, but the amount of deposits that they receive. This is the multiplication of the probability that they are chosen by the amount of deposits that they receive conditional on being chosen. Multiplying the probability that bank $j$ is chosen ($Pr_{r_j}^d$) with the amount of deposits held at bank $j$ if it is chosen ($d_j$), one obtains:

$$d_j Pr_{r_j}^d = (1 - \alpha^d)(1 + i_j^d)^{\theta^d-1}(p_d^j)^{\theta^d-\epsilon_d^j}(p^d)^{\epsilon_d^j-1}.$$

4 After adding a logarithm to the utility function, which does not alter the maximization problem.
5 As mentioned earlier, this stochastic component can be due to several things: geographic variation, switching costs, recommendations of family or friends, etc.
I interpret \((1 - \alpha^d)(1 + i^d)\theta^{d-1}(\rho^d)^{\theta^d-1}Y^d\) as aggregate deposits and denote it with \(D\). Even though this quantity varies with the policy rate, here I will keep \(D\) fixed and ignore its dependence on the policy rate. I do this in order to preserve the partial equilibrium nature of the model in Ulate (2019), which assumes that the banks optimize in response to changes in the policy rate but the aggregate amount of loans and deposits remains fixed. Additionally, I interpret \(d_j Pr^d_j\) as the amount deposited to bank \(j\) once the whole population of savers is taken into account, and denote this by \(D_j\). Then:

\[
D_j = \left(\frac{1 + i^d_j}{1 + i^d}\right)^{\theta^d-1} \left(\frac{p^d_j}{\rho^d}\right)^{\theta^d \varepsilon^d} D.
\] (5)

This means that deposit supply for bank \(j\) has two distinct elasticity margins.

Using equation (5) and the definition of \(p^d_j\), the elasticity of deposit supply with respect to the gross deposit rate can be written as:

\[
\gamma^d_j \equiv \frac{\partial D_i}{\partial (1 + i^d_j)} \frac{1 + i^d_j}{D_j} = s^d_j (\varepsilon^d - 1) + (1 - s^d_j) (\theta^d - 1) = (\theta^d - 1) - s^d_j (\theta^d - \varepsilon^d),
\] (6)

where \(s^d_j \equiv ((1 - \alpha^d)(1 + i^d_j)^{\theta^d-1})/(\alpha^d + (1 - \alpha^d)(1 + i^d_j)^{\theta^d-1})\), \(0 \leq s^d_j \leq 1\), and \(\partial s^d_j / \partial i^d_j > 0\). When the deposit rate charged by bank \(j\) is high, the weight \(s^d_j\) is high, and the elasticity is driven towards \(\varepsilon^d - 1\). Conversely, when \(i^d_j\) is low, the weight \(s^d_j\) is low, and the elasticity is driven towards \(\theta^d - 1\). To the extent that \(\theta^d > \varepsilon^d\), increasing the policy rate (which will increase the deposit rate of all banks) decreases the elasticity and leads to higher markups. This implies that the pass-through from the policy rate to the deposit rate is smaller than one.

When \(i^d_j\) is low, the price of consumption of bank \(j\) \((p^d_j\) tends to 1. Hence, the second parenthesis in (5) plays a smaller role and the main elasticity left in the expression for deposit supply is \(\theta^d - 1\). Intuitively, when the deposit rate is low, it is not an important factor in how customers substitute between banks, and hence the main elasticity that determines deposit supply is the intertemporal elasticity of substitution. In contrast, a high deposit rate \(i^d_j\) tilts the price of consumption for bank \(j\) towards \((1 + i^d_j)^{-1}\), allowing the second parenthesis in (5) to be combined with the first, with an elasticity of \(\varepsilon^d - 1\). Intuitively, with a high deposit

---

6There are several margins besides the ones considered here which can affect deposit supply. Those margins might even dominate the influence of the policy rate. That is why in this paper I choose to abstract from analyzing changes in aggregate deposit supply and focus instead on the allocation of such aggregate supply.

7In the limit of a continuous time model, the elasticity of substitution between “today” and “tomorrow” should be high, approaching infinity. In contrast, the elasticity of substitution between banks will remain bounded because the switching costs operate across several periods. It is then natural to expect \(\theta^d > \varepsilon^d\).
rate (and $\theta^d > 1$), most consumption happens tomorrow, making the I.E.S. irrelevant and turning $\varepsilon^d - 1$ into the crucial elasticity governing deposit supply.

### 3.2 The Problem of the Borrowers

The problem of the borrower is somewhat related to the one of the saver, but it is slightly different. A borrower has CES preferences between today and tomorrow:

$$U(C_0, C_1) = \left[ (\alpha^l)^{\frac{1}{\theta^l}} C_0^{\frac{\theta^l}{\theta^l-1}} + (1 - \alpha^l)^{\frac{1}{\theta^l}} C_1^{\frac{\theta^l}{\theta^l-1}} \right]^{\frac{\theta^l}{\theta^l-1}}.$$

where $C_0$ is consumption today, $C_1$ is consumption tomorrow, $\alpha^l$ is the importance of consumption today, and $\theta^l$ is the elasticity of substitution between consumption today and consumption tomorrow. In contrast with the saver, the borrower only has income $\mathbf{Y}^d$ tomorrow. He needs to borrow in order to consume today. He can borrow from a continuum of banks between 0 and 1. The budget constraint conditional on the choice of bank $j$ can be expressed as:

$$(1 + i^d_j)C_0 + C_1 = \mathbf{Y}^d.$$

The solution to this problem is:

$$C_0 = \alpha^l \left( \frac{1 + i^d_j}{p_j^l} \right)^{-\theta^l} \frac{\mathbf{Y}^d}{p_j^l}, \quad C_1 = (1 - \alpha^l) \left( \frac{1}{p_j^l} \right)^{-\theta^l} \frac{\mathbf{Y}^d}{p_j^l},$$

where

$$p_j^l \equiv \left( \alpha^l(1 + i^d_j)^{1-\theta^l} + 1 - \alpha^l \right)^{1-\theta^l}$$

is the price index of aggregate consumption for a borrower that chooses bank $j$.

Taking the same stochastic utility approach as in the case of the saver, the probability for a consumer of choosing bank $j$ is given by: $Pr^d_j = (p_j^l/p^l)^{1-\varepsilon^l}$, where $p^l$ is the usual CES price index: $p^l = \left( \int_0^1 (p^l_r)^{1-\varepsilon^l} dr \right)^{1-\varepsilon^l}$. Multiplying the amount borrowed from bank $j$ if it is chosen ($B_j$) by this probability, one obtains:

$$B_j Pr^d_j = \alpha^l(1 + i^d_j)^{-\theta^l}(p_j^l)^{\theta^l-\varepsilon^l}(p^l)^{\varepsilon^l-1}\mathbf{Y}^d.$$

Interpret $\alpha^l(1 + i^d)^{-\theta^l}(p^l)^{\theta^l-1}\mathbf{Y}^d$ as aggregate borrowing and denote it as $\mathbf{L}$. Additionally, interpret $B_j Pr^d_j$ as the amount borrowed from each bank $j$ once the whole population of
borrowers is taken into account, and denote this by $L_j$. Then:

$$L_j = \left( \frac{1 + i^l_j}{1 + i^l_j} \right)^{-\theta^l} \left( \frac{p_j^l}{p^d} \right)^{\theta^d - \epsilon^l} L.$$  \hfill (7)

The interpretation of this equation is similar to the one of equation (5). The elasticity of loan demand with respect to the gross loan rate is:

$$\gamma^l_j \equiv \frac{\partial L_j}{\partial (1 + i^l_j)} \frac{1 + i^l_j}{L_j} = -s^l_j \varepsilon^l - (1 - s^l_j) \theta^l = -s^l_j (\theta^l - \varepsilon^l),$$  \hfill (8)

where $s^l_j \equiv (\alpha^l (1 + i^l_j)^{-\theta^l}) / (\alpha^l (1 + i^l_j)^{-\theta^l} + 1 - \alpha^l), 0 \leq s^l_j \leq 1,$ and $\frac{\partial s^l_j}{\partial i^l_j} < 0.$

### 3.3 The Problem of the Banks

The setup of the banking problem is similar to the one in Section 2, but here I also introduce exogenous costs of issuing loans and deposits ($\mu^l$ and $\mu^d$). Banks choose the interest rate they charge on loans $i^l_j$, the amount they lend, the interest rate they pay on deposits $i^d_j$, the amount of deposits they take, and the amount of reserves they hold at the central bank, which earns the policy rate $i$, subject to several constraints. The maximization problem that the individual bank $j$ faces is therefore the following:

$$\max_{i^l_j, L_j, i^d_j, D_j, H_j} \quad (1 + i^l_j - \mu^l)L_j + (1 + i)H_j - (1 + i^d_j + \mu^d)D_j$$

s.t.

$$L_j = \left( \frac{1 + i^l_j}{1 + i^l_j} \right)^{-\theta^l} \left( \frac{p_j^l}{p^d} \right)^{\theta^d - \epsilon^l} L$$

$$D_j = \begin{cases} 
(1 + s^d_j)^{\theta^d - 1} \left( \frac{p_j^d}{p^d} \right)^{\theta^d - \epsilon^d} \quad & \text{if } i^d_j \geq 0 \\
0 & \text{if } i^d_j < 0 
\end{cases}$$

$$L_j + H_j = F_j + D_j$$

$$H_j \geq 0,$$

where $\mu^l$ is the cost of issuing loans and $\mu^d$ is the cost of issuing deposits.

In Regime 1, where the banks can solve their problem unconstrained by the ZLB on deposits and optimally hold positive reserves, the F.O.C.’s w.r.t. the gross loan rate and the
gross deposit rate are:

\[
0 = L_j + [(1 + i^l_j) - (1 + i) - \mu^l] \frac{\partial L_j}{\partial (1 + i^l_j)}
\]

\[
0 = -D_j + [(1 + i) - (1 + i^d_j) - \mu^d] \frac{\partial D_j}{\partial (1 + i^d_j)}.
\]

Using the elasticities provided in equations (6) and (8) these equations can be simplified to:

\[
1 + i^l_j = \frac{\gamma^l_j}{\gamma^l_j + 1} (1 + i + \mu^l),
\]

\[
1 + i^d_j = \frac{\gamma^d_j}{\gamma^d_j + 1} (1 + i - \mu^d).
\]

Since the elasticities \(\gamma^l_j\) and \(\gamma^d_j\) contain \(i^l_j\) and \(i^d_j\) respectively, these equations don’t provide a closed-form solution for the loan rate and the deposit rate, but they can be solved numerically. Nevertheless, the previous equations are still very useful, since they clarify that the gross loan rate is set as a markup (since \(\gamma^l_j < -1\)) on the gross policy rate and the gross deposit rate is set as a markdown (since \(\gamma^d_j > 0\)) on the gross policy rate. Since all banks are identical, they all charge the same loan rate and pay the same deposit rate (denoted by \(i^l\) and \(i^d\)). Return on equity for banks is then given by:

\[
\frac{F'}{F} - 1 = i + (i^l - i - \mu^l) \frac{L}{F} + (i - i^d - \mu^d) \frac{D}{F}.
\]

In Regime 2 banks pay a zero rate on deposits and obtain a fixed amount of deposits \(D\), choose the interest rate they charge on loans \(i^l_j\), the amount they lend, and the amount of reserves they hold in the central bank. The solution for the loan rate is exactly the same one as in Regime 1. Return on equity has the same expression as in Regime 1 after setting \(i^d = 0\). The solution for Regime 3 is a bit complicated, but works very similarly than Regime 3 in the original static model of Ulate (2019).

If I assume parameter values \(\alpha^d = \alpha^l = 0.9\), \(\varepsilon^d = \varepsilon^l = 10\), \(\theta^l = \theta^d = 100\), \(\mu^l = 0.8\%\), \(\mu^d = -0.6\%\), \(D/F = 9\) and \(L/F = 10\), then the behavior of rates is the one illustrated in Figure 2 and the behavior of ROE is the one illustrated by the blue line in Figure 4. Importantly, the model exhibits non-unitary pass-through similar to the one in the data. While the parameter values that I assume (in order to obtain a pass-through that can match the stylized facts described in the introduction) are not carefully calibrated, this setup illustrates the fact that models with a non-unitary pass-through can still feature a behavior of bank ROE that is similar to the one in Ulate (2019).
4 “Liquidity and Bonds” Model

The second extended model relies on different mechanisms to generate a non-unitary pass-through in the loan rate and the deposit rate. On the loan side the mechanism is exactly the same as in the previous model. Consequently, the problem of the borrowers is not described here. The problem of the savers is different, since the intertemporal elasticity of substitution is once again assumed to be unitary, but agents can now save in cash, bonds, or deposits with a continuum of banks. The next subsection describes the problem of the saver, and the following one describes the bank’s problem.

4.1 The Problem of the Savers

I assume that there is an individual consumer that lives for two periods, denoted 0 and 1. This consumer has a total income of \(Y^d\) in the first period and he can consume in both periods. To consume in period 0 is easy for this consumer, it can be done directly. However, to consume in period 1 the consumer must save some of his current income \(Y^d\). He can save in three ways: through one of a continuum of banks between zero and one (indexed with \(j\)), in cash (which offers a nominal return of zero percent), or in bonds that pay a gross return of \((1+i)\).

The decision process of this consumer happens in two stages. In the first stage, the consumer decides which bank he wants to save with, and in the second stage he chooses the amounts he wants to allocate to cash, deposits, and bonds. First, I will describe the problem of a consumer that has already chosen bank \(j\), and then I will describe the way that the bank choice is made. I assume that the direct utility function of the consumer conditional on his choice of bank \(j\) is given by:

\[
U(C_0, C_1, L_j) = \ln(C_0) + \beta \ln(C_1) + \gamma \ln(L_j),
\]

where \(\beta\) is the discount factor between periods, \(\gamma\) is the importance of liquidity in utility, and \(C_t\) is consumption in period \(t\). Additionally, \(L_j\) represents liquidity services, which are the following combination of deposits in bank \(j\) and cash:

\[
L_j = \left( \left( \alpha^d \right)^{\frac{1}{\theta_d}} d_j^{\frac{\theta_d - 1}{\theta_d}} + (1 - \alpha^d)^{\frac{1}{\theta_d}} M_j^{\frac{\theta_d - 1}{\theta_d}} \right)^{\frac{\theta_d}{\theta_d - 1}},
\]

where \(\alpha^d\) is the importance of deposits in liquidity provision, \(\theta^d\) is the elasticity of substitution between cash and deposits in liquidity provision, \(d_j\) are deposits at bank \(j\), and \(M_j\) is the amount of cash held conditional on the choice of bank \(j\).

The first and second period budget constraints of the saver (again, conditional on the
choice of bank $j$ are:

$$
P_0 C_0 = P_0 Y^d - d_j - M_j - B_j
$$

$$
P_1 C_1 = (1 + i^d_j)d_j + (1 + i)B_j + M_j,
$$

where $1 + i^d_j$ is the gross deposit rate paid between periods 0 and 1 by bank $j$ (which is known by the consumer with certainty), $B_j$ is the amount of bonds held conditional on the choice of bank $j$, $i$ is the policy rate (which is assumed to be the return on bonds), and $P_t$ is the price index in period $t$. The aggregate budget constraint can then be expressed as:

$$
C_0 = Y^d - \left( \frac{1}{1 + i} \right) P_0 C_1 - \frac{i - i^d_j}{1 + i} M_j - \frac{1}{1 + i} B_j P_0
$$

The solution to the saver's problem conditional on the choice of bank is:

$$
C_0 = \frac{Y^d}{1 + \beta + \gamma}, \quad C_1 = \frac{\beta(1 + r)}{1 + \beta + \gamma} Y^d,
$$

$$
L_j = \frac{\gamma(1 + i) Y^d}{1 + \beta + \gamma p^d_j}, \quad d_j = \alpha^d \left( \frac{i - i^d_j}{p^d_j} \right)^{1 - \theta^d} L_j,
$$

where $P_0$ has been normalized to one, $1 + r \equiv (1 + i)\frac{P_1}{P_0}$, and the price of liquidity $p^d_j$ is given by

$$
p^d_j = \left[ \alpha^d (i - i^d_j)^{1 - \theta^d} + (1 - \alpha^d)i^{1 - \theta^d} \right] \frac{1}{1 - \theta^d}.
$$

With these quantities, the indirect utility function conditional on borrowing from bank $j$ can be expressed as:

$$
v^d_j = (1 + \beta + \gamma)(\ln Y^d - \ln(1 + \beta + \gamma)) + \beta \ln(\beta) + \gamma \ln(\gamma) + \beta \ln(1 + r) + \gamma \ln(1 + i) - \gamma \ln(p^d_j)
$$

Then, as in Anderson et al. (1988), assume that the first stage (the bank choice stage), is described by a stochastic utility approach:

$$
V^d_j = v^d_j + \mu^d \epsilon^d_j,
$$

where $\mu^d$ is a positive constant and $\epsilon^d_j$ is a random variable with zero mean and unit variance. Assuming that the $\epsilon^d_j$ random variables are independently and identically distributed with
type-one extreme value distribution, the probability for a consumer of choosing bank \( j \) is:

\[
Pr_j^d = Pr(V_j^d = \max_r V_r^d) = \frac{e^{v_j^d/\mu^d}}{\int_0^1 e^{v_r^d/\mu^d} dr} = \frac{(p_j^d)^{-\gamma/\mu^d}}{\int_0^1 (p_r^d)^{-\gamma/\mu^d} dr}.
\]

Substituting \(-\gamma/\mu^d\) for \(1 - \varepsilon^d\), the previous expression can be rewritten as:

\[
Pr_j^d = \left(\frac{p_j^d}{p^d}\right)^{1-\varepsilon^d}
\]

where \(p^d\) is the aggregate price of liquidity defined in the usual way. Multiplying \(d_j\) by this probability and simplifying one obtains:

\[
d_j Pr_j^d = \alpha^d \gamma (1 + i) \frac{\gamma^d}{\gamma^d + \gamma^d + \gamma^d} \left(\frac{i - i_j^d}{p^d}\right)^{-\gamma^d} \left(\frac{i - i_j^d}{i - i^d}\right)^{-\gamma^d} \left(\frac{p_j^d}{p^d}\right)^{\gamma^d - \varepsilon^d}.
\]

Interpret \(\alpha^d \gamma (1 + i) \frac{\gamma^d}{\gamma^d + \gamma^d + \gamma^d} \left(\frac{i - i_j^d}{p^d}\right)^{-\gamma^d} \left(\frac{i - i_j^d}{i - i^d}\right)^{-\gamma^d} \left(\frac{p_j^d}{p^d}\right)^{\gamma^d - \varepsilon^d}\) as aggregate deposits and denote them with \(D\). Additionally, interpret \(d_j Pr(j)\) as the amount deposited in bank \( j \) once the whole population of consumers is taken into account, and denote this by \(D_j\). Then:

\[
D_j = \left(\frac{i - i_j^d}{i - i^d}\right)^{-\gamma^d} \left(\frac{p_j^d}{p^d}\right)^{\gamma^d - \varepsilon^d} D,
\]

This is related to equation (5) but it is different in several aspects. First, the exponent of the first term is \(-\theta^d\) instead of \(\theta^d - 1\). Second, the quantity inside the first parenthesis is a ratio of spreads \((i - i_j^d)\) instead of a ratio of gross interest rates (because now the customers have a bigger selection of saving instruments). Third, the definition of \(p_j^d\) is different in this context.

4.2 The Problem of the Banks

The setup of the banking problem is exactly the same as in Section 3.3, with a single change to make deposit supply follow (9) instead of (5). The maximization problem that individual bank \( j \) faces is therefore the following:

\[
\max_{i_j^d, L_j, i_j^d, D_j, H_j} (1 + i_j^d - \mu^d) L_j + (1 + i) H_j - (1 + i_j^d + \mu^d) D_j
\]

s.t.
\[
L_j = \left( \frac{1 + i_j}{1 + i_l} \right)^{-\theta_l} \left( \frac{p_l}{p_l'} \right)^{\theta_l - \varepsilon_l} L
\]
\[
D_j = \begin{cases} 
\left( \frac{i - i_d}{i - i_l} \right)^{-\theta_d} \left( \frac{p_d}{p_d'} \right)^{\theta_d - \varepsilon_d} D & \text{if } i_d \geq 0 \\
0 & \text{if } i_d < 0
\end{cases}
\]
\[
L_j + H_j = F_j + D_j
\]
\[
H_j \geq 0,
\]

The F.O.C for the loan rate is exactly the same as in Section 3.3. Meanwhile, the derivative of deposit supply w.r.t. \( i_d j \) is:
\[
\frac{\partial D_j}{\partial i_d j} = \theta_d \frac{D_j}{i - i_d j} - (\theta_d - \varepsilon_d) \frac{D_j}{p_d j} \alpha_d \left( \frac{i - i_d j}{p_d j} \right)^{-\theta_d}.
\]

The F.O.C. w.r.t. \( i_d j \) is the following:
\[
0 = -D_j + (i - i_d j - \mu_d) \frac{\partial D_j}{\partial i_d j}.
\]

Combining the previous two equations, one obtains:
\[
0 = \alpha_d (i - i_d j)^2 - \theta_d(1 - \varepsilon_d) + (i - i_d j)(1 - \alpha_d) i^{1-\theta_d}(1 - \theta_d) + \mu_d \varepsilon_d (1 - \alpha_d) i^{1-\theta_d} + \mu_d \varepsilon_d \alpha_d (i - i_d j)^{1-\theta_d}.
\]

As in the previous extended model, this equation cannot be solved explicitly for \( i_d j \), but it can be solved numerically. Return on equity for banks is described by the same expression as in the first extended model, and Regime 2 and 3 work in a similar way as well.

If I assume parameter values \( \alpha_l = 0.9, \varepsilon_l = 10, \theta_l = 100 \) and \( \mu_l = 0.8\% \) on the loan side, and \( \alpha_d = 0.5, \varepsilon_d = 2, \theta_d = 0.5 \) and \( \mu_d = 0.1\% \) on the deposit side, as well as \( D/F = 9 \) and \( L/F = 10 \), then the behavior of rates is the one illustrated in Figure 3 and the behavior of ROE is the one illustrated by the red line in Figure 4. As in the case of the “High Intertemporal Substitution” model, this second model also exhibits non-unitary pass-through similar to that in the data. The parameter values that I assume on the loan side are the same as the ones assumed in the “High Intertemporal Substitution” model. On the deposit side, \( \alpha_d = 0.5 \) indicates that deposits and cash have the same importance in the liquidity aggregator, \( \theta_d = 0.5 \) indicates that deposits and cash are not very substitutable, and \( \varepsilon_d = 2 \) indicates that banks have substantial monopoly power.
While this model ends up delivering a behavior of rates similar to that in the “High Intertemporal Substitution” model, it relies on a completely different mechanism to deliver non-unitary deposit pass-through. This can be beneficial to researchers that want to include cash and bonds in their general equilibrium models for alternative reasons. Additionally, the saving parameter values are more realistic in this extension.

5 Conclusion

This paper proposes static and partial equilibrium models of the banking sector in order to study the pass-through of the policy rate to the loan rate and the deposit rate. First, the paper discusses the partial equilibrium model of Ulate (2019). This model is useful to convey intuition and to study negative nominal interest rates, but it features a complete pass-through of the policy rate to loans and deposit rates in “normal territory”. This complete pass-through is not consistent with stylized facts indicating that the pass-through of the policy rate to loan and deposit rates is between 0.5 and 0.8 during normal times.

Next, the paper modifies the static framework of Ulate (2019) and proposes two models which can match the aforementioned stylized facts while remaining parsimonious. Importantly, the proposed models do not rely on having large banks to obtain a realistic pass-through, as they can deliver a non-unitary pass-through even with a continuum of banks.

The first model relies on a CES utility function between today and tomorrow, a sequential choice of bank and saving (or borrowing) amounts, and differentiation between banks. For borrowers, an intertemporal elasticity of substitution greater than one implies that they want to borrow a small amount when rates are high. This gives lenders “less monopoly power” and makes them charge a smaller loan spread when rates are high. In contrast, banks charge a higher deposit spread when rates are high.

In the second model, agents can save using cash, deposits in a continuum of banks, or bonds. Additionally, cash and bonds provide liquidity services through a CES aggregator. Savers must first choose a bank and only then choose their allocations (amount of deposits, cash, or bonds). When rates are high, the return differential between bonds and cash is high, making deposits valuable and allowing banks to charge a high deposit spread.

Overall, the extended models provide a parsimonious way of capturing non-unitary pass-through in normal territory, while also providing realistic pass-through in negative territory. Additionally, they do not require a small number of banks and hence sidestep the associated complication of determining the evolution of the number of banks. Moreover, the extended models have similar implications for return on equity as the static model of Ulate (2019), and they suggest that having a non-unitary pass-through in normal territory does not modify substantially the analysis under negative rates.
References


