Funds Transfer Pricing, Liquidity Premium, and Market Structure

Ted Lindblom and Viktor Elliot
School of Business, Economics and Law at University of Gothenburg
E-mail: ted.lindblom@handels.gu.se
viktor.elliot@handels.gu.se

Abstract
Funds transfer pricing (FTP) is widely acknowledged as an important part of banks’ asset and liability management (ALM). This paper makes two contributions to the existing FTP-theory. First, we analyze the implications of adding a liquidity premium to the FTP. Second, since a majority of the existing banking markets are dominated by a limited number of large banks we also cover the case of kinked demand and supply curves on oligopolistic and oligopsonistic competition, respectively.

Keywords Funds Transfer Pricing, FTP, Oligopoly, Oligopsony, Banks

1. Introduction
Funds Transfer Pricing (FTP) plays a fundamental role in the asset and liability management (ALM) of the modern banking firm. The FTP is used to measure, determine and evaluate the profitability of all outstanding loans (L) on the bank’s asset side as well as the profitability of each deposit and savings product (D) on its liability side (Dermine, 2009). The FTP is also essential in the decentralized bank’s internal assessments of managerial performance, customer relationship profitability and, branch efficiency. When based on market rates set on efficient money and capital markets, the FTP provides decision-makers in the bank with the actual opportunity costs of funds (κi) at which a granted loan li (i = 1,2..n) could be financed, and the corresponding cost (κd) at which an attracted deposit (and savings product) di (j = 1,2..m) could be invested, on the financial market. As long as the bank has access to efficient financial markets at negligible or very low transaction costs, the asset and liability decisions can thus be made separately and independently of one another. The ALM separation has implications for the bank’s delegation of responsibility and, ultimately, how the modern banking firm is organized for efficiently managing exposures to risks, enhancing value creation and complying with regulatory requirements.

The use of FTP in banks has policy implications for the functioning and stability of the entire financial system (Dermine, 2011; Grant, 2011; Elliot, 2015). The recent financial crisis brought attention to the fact that market rates do not account for the exposure to liquidity risk of individual banking firms (King, 2013; Dietrich, Hess and Wanzenried, 2014; Bonner, van Lelyveld and Zymek, 2015; Bologna, 2015). As emphasized by Cadamagnani et al. (2015), this implies an inadequate or lacking pricing of liquidity risk in the market-based FTP, which is likely to have contributed to excessive maturity mismatching prior to the crisis. Regulators quickly acknowledged this problem and are now expecting the banking firm to account for, and report on, its liquidity risk management practice at a much more detailed level (cf. BIS, 2008a,b; 2010; 2013a,b). The importance of FTP for efficient liquidity risk management is evident in these regulatory texts (see specifically BIS, 2008a,b; EBA 2010). In order to accurately mirror the opportunity cost of each specific source of funds, regulators and many scholars today advocate that banks should “add-on” a liquidity premium when using the market-based FTP (cf., Dermine, 2011; Grant, 2011; Tumasyan, 2012; van Deventer, Imai and Mesler, 2013; Elliot, 2015).
The implementation of a liquidity premium has implications for the ALM of banks. In a decentralized bank, the premium should incentivize branch managers to consider also the cost of liquidity risk exposures in their loans and deposits decisions. To accurately price the liquidity risk, the liquidity premium should reflect the opportunity cost of such exposures, which according to Grant (2011:30) can be derived by: "converting fixed-rate borrowing costs to floating-rate borrowing costs through an internal swap transaction and observing the spread over the reference rate, which is depicted from the swap curve". The spread is defined as the term liquidity premium and is likely to increase with the fund's maturity. When added on the market-based FTP, the bank’s internal prices on funds will deviate from the prices on the financial market. By definition, the adjusted FTP will equal the bank’s marginal (opportunity) cost of borrowing funds on the market, but not necessarily the marginal gain or benefit for the bank of being a net provider of funds to the market. Hence, it is not evident that the adjusted FTP will make the decisions at the branch level fully aligned with the overall objective of the bank.

The alignment between branch decisions and the overall objective of the bank is even less evident if it operates on oligopolistic and/or oligopsonistic markets characterized by price stickiness. Already in the early 1970s, Monti (1972:435) observed that: "markets for bank assets and liabilities are often oligopolistic in nature". In their extensive study of bank regulation and supervision, Barth, Caprio and Levine (2013:115) find that the degree of bank concentration is still at a high level in most countries, with the median share of bank assets among the top five banks in each country being 73 percent. There are also frequent observations of price rigidity (cf. Jacksson, 1997; 2014; Peltzman 2000; De Graeve, De Jonghe and Vennet, 2007), which may be caused by so-called Sweezy (1939) competition. Under Sweezy competition, bank (branch) managers are reluctant to change their interest rates on loans \(r_i\) and deposits \(r_d\) unless major competitors do so. Each branch manager knows that the bank’s competitors would retaliate if lending rates were cut and/or deposit rates were raised, but not respond if the rates were changed in the opposite direction. Consequently, the branch managers are facing "kinked" demand and supply curves for loans and deposits, respectively.

This paper aims at analyzing and demonstrating how the addition of a liquidity premium to a bank’s FTP will affect its optimal volumes of loans and deposits under different market conditions. In the current FTP theory, the liquidity premium and its potential effects are not covered explicitly—let alone the cases of oligopoly and/or oligopsony where the price rigidity of banks is high.

Our analysis suggests that it is optimal for a bank to adjust its market-based FTP by adding-on a liquidity premium when selling and purchasing funds internally, as long as the bank is a net borrower of these funds on the financial market and does not operate under Sweezy-like competition on either the loan market or the deposit market. Assuming that the bank is a net funds provider on the financial market, and particularly under high price rigidity on the loan and/or deposit market, it would instead be optimal for the bank to differentiate the liquidity premium. In the case of oligopoly on the loan market, this could mean that a premium should be added-on the FTP when purchasing funds only. In the case of oligopsony on the deposit market, the liquidity premium may only be motivated on funds sold. If both
these markets are characterized by price stickiness, the add-on of a liquidity premium may be too costly for the bank and not motivated at all.

The paper is organised as follows. In the next section, we briefly recapitulate the current FTP theory, after which in section 3 we extend the theory by illustrating what will happen when a liquidity premium is added-on to the FTP. In Section 4, we then analyse the cases of oligopoly and oligopsony under the assumption of kinked loan demand and deposit supply curves, respectively. Section 5 concludes the paper.

2. Review of current funds transfer pricing theory

The FTP theory builds on the separation theorem (Klein, 1971; Monti, 1972) and implies that maturity imbalances between assets and liabilities in the bank’s balance sheet are settled on the financial market. In the plain vanilla setting, however, the bank has no access to efficient financial markets and its outstanding loans ($L_t$) and deposits available for lending ($D_t$) have the same maturity $t$. The bank operates two branches of which one is loan-oriented (Branch $L$) and possesses monopoly power on the loan market. Accordingly, the other branch (Branch $D$) is deposit-oriented with monopsony power on the deposit market. Without access to an efficient financial market, Branch $L$ must finance all lending that exceeds its deposit volume by purchasing “surplus” funds from Branch $D$. This suggests that $L_t \leq D_t$. In turn, Branch $D$ must sell all deposits that exceeds its level of lending to Branch $L$ suggesting that $L_t \geq D_t$. These constraints imply that $L_t = D_t$.

In this setting, the profit of the bank ($\pi$) is maximized at the volume ($V_t^*$) where its net marginal revenue of lending ($NMR(L_t)$) equals its marginal cost of deposits ($MC(D_t)$) given that $\frac{\delta NMR(L_t)}{\delta L_t} < \frac{\delta MC(D_t)}{\delta D_t}$ and, thus, the bank’s profit function shown in Equation 1 is strictly concave around $V_t^*$.

$$\max_{L_t,D_t} \pi = \int_0^{V_t^*} (\tau(L_t) - \omega(L_t)) dV_t - \int_0^{V_t^*} (\tau(D_t) + \omega(D_t)) dV_t - F,$$  \hfill (1)

subject to: $L_t = D_t$,

$$\delta \pi \over \delta L_t = \tau(L_t) - \omega(L_t) - \lambda_t = 0,$$  \hfill (2)

$$\delta \pi \over \delta D_t = - (\tau(D_t) + \omega(D_t)) + \lambda_t = 0,$$  \hfill (3)

$$\tau(L_t) - \omega(L_t) = \tau(D_t) + \omega(D_t),$$  \hfill (4)

$$NMR(L_t) = MC(D_t),$$

where

$\pi$ = Profit of the bank,
$L_t$ = Volume of loans with maturity $t$,
$D_t$ = Volume of deposits with maturity $t$,
$V_t^*$ = Optimal volume of loan/deposits with maturity $t$,
$\tau(L_t)$ = Marginal interest revenue from loans with maturity $t$,
$\tau(D_t)$ = Marginal interest cost to provide deposits with maturity $t$,
$\omega(L_t)$ = Marginal operational cost of loans with maturity $t$. 

3
\(\omega(D_t) = \text{Marginal operational cost of deposits with maturity } t,\)
\(F = \text{Fixed costs,}\)
\(\lambda_t = \text{Lagrange multiplier (reflects the opportunity cost of loans/deposits with maturity } t),\)
\(NMR(L_t) = \text{Net marginal revenue of loans with maturity } t,\)
\(MC(D_t) = \text{Marginal cost of deposits with maturity } t.\)

The idea of an internal market is fundamental and generic to literature on transfer pricing (cf., Hirshleifer, 1956) and the basic principles are the same for banks. In Equation 1, \(NMR(L_t)\) reflects the “internal demand” for funds with maturity \(t\) in the bank and is derived by deducting the marginal operational cost of lending \(\omega(L_t)\), for attracting and retaining loan customers from the bank’s marginal interest revenue \(\tau(L_t)\). Accordingly, \(MC(D_t)\) expresses the bank’s “internal supply” of funds with maturity \(t\) and equals the sum of the bank’s marginal interest cost \(\tau(D_t)\) and its incurred marginal operational costs to attract deposits \(\omega(D_t)\).

In the modern bank, the treasury department (henceforward referred to as treasury) executes the process of matching loans and deposits to financial securities of the same maturity and term structure on the markets (cf. Dermine, 2009; van Deventer, Imai and Mesler, 2013). Treasury is accountable for managing the bank’s overall exposures to liquidity and interest rate risk. It acts as a market maker within the bank to which the branches turn for settling their imbalances of loans and deposits of different maturities \(t\) against their current internal market price \(FTP_t\). Consequently, the branches are left with exposures to the credit and/or funding risks only. The \(FTP_t\) used by treasury will incur additional costs for Branch \(L\) and give rise to the corresponding additional revenue for Branch \(D\). Provided that the current \(FTP_t\) reflects the opportunity cost of each loan \(l_{i,t} (k_{i,t}^t)\) as well as each deposit \(d_{j,t} (k_{j,t}^t)\) with maturity \(t\), the true contribution of each branch to the bank’s overall profit can be determined. This paves the way for performance measurement and decentralized decision-making in the bank.

In the naive case that treasury has continuous access to accurate information about the internal demand and supply conditions at the branch level, treasury may derive and set the optimal \(FTP_t^o\) for funds with maturity \(t\) centrally. When the bank has no access to an efficient financial market and all funds have the same maturity \(t\), the optimal \(FTP_t^o\) is equal to the interest rate at which the internal demand for funds \((NMR(L_t))\) coincides with the internal supply of funds \((MC(D_t))\). Hence, \(FTP_t^o\) reflects the opportunity cost of loans and deposits with maturity \(t\). Thereby, each branch can derive its optimal loan volume \((L_t^o)\) as well as its optimal deposit volume \((D_t^o)\) separately. This is demonstrated in Equations 2 and 3 in which the bank’s profit function is broken down into the strictly concave profit functions for the loan-oriented branch \((\pi_L)\) and the deposit-oriented branch \((\pi_D)\) around \(L_t^o\) and \(D_t^o\), respectively, under the assumption that the latter only takes deposits and the former only grants loans. Such extreme specialization of the branches is of course uncommon in practice, but if treasury is assumed to either break even or incur no additional costs it highlights that the overall profit for the bank is the sum of the profits of its parts (branches): \(\pi = \pi_L + \pi_D.\)

\[
\max_{L_t} \pi_L = \int (\tau(L_t) - \omega(L_t))dL_t - FTP_t^o \cdot L_t, \tag{2}
\]
\[
\leq\Rightarrow \delta \pi_L = \tau(L_t) - \omega(L_t) - FTP_t^o = 0 \Rightarrow FTP_t^o = \tau(L_t^o) - \omega(L_t^o) = NMR(L_t^o).
\]
\[ \max_{D_t} \pi_D = FTP^o_t \cdot D_t - \int (\tau(D_t) + \omega(D_t)) dD_t, \]  

\[ \Rightarrow \delta \pi_D = FTP^o_t - \tau(D_t) - \omega(D_t) = 0 \Rightarrow FTP^o_t = \tau(D_t^o) + \omega(D_t^o) = MC(D_t^o), \]

where
\( \pi_D \) = Profit of the lending-oriented branch,
\( \pi_L \) = Profit of the lending-oriented branch,
\( FTP^o_t \) = The optimal transfer price of funds with maturity \( t \) without market access,
\( L_t^o \) = Optimal volume of loans with maturity \( t \) without market access,
\( D_t^o \) = Optimal volume of deposits with maturity \( t \) without market access.

The full potential of decentralized decision-making in banks is reached when there is low-cost access to efficient financial markets. Then the volume constraint, \( L_t = D_t \), can be relaxed and, moreover, treasury no longer requires knowledge about the internal demand and supply conditions at the branch level. Under such conditions, the separation theorem implies that the bank’s branches should make the decisions about loan and deposit volumes separately (at least in the short run) regardless of the volume at which \( NMR(L_t) \) equals \( MC(D_t) \). Given that \( NMR(L_t) \) is linearly decreasing with loan volume and \( MC(D_t) \) is linearly increasing with deposit volume it is optimal for the bank that treasury settles all imbalances between its outstanding loans and deposits to the relevant market interest rate (see Figure 1). This rate gives the market-based \( FTP^M_t \).

The shaded areas in the diagrams in Figure 1 show the increase in the bank’s overall profit of adopting \( FTP^M_t \) instead of \( FTP^o_t \). With low-cost access to efficient financial markets, \( FTP^M_t \) reflects the opportunity cost of maturity \( t \) funds and, thus, both the relevant \( k^l_t \) for loans as well as the relevant \( k^d_t \) for deposits of this maturity. The diagram on the left-hand side illustrates a case when the actual market rate of a security with the corresponding maturity is higher than the optimal \( FTP^o_t \) for the bank without access to efficient financial markets, i.e. the market rate is higher than the bank’s \( NMR(L_t) \) and \( MC(D_t) \) at the volume where \( L_t = D_t \). Thus, the bank will be a net provider of funds with this maturity on the financial market. The diagram to the right shows the opposite situation. Here, the bank is instead a net funds borrower on the market. Both diagrams display how separate lending and funding decisions on the branch level will maximize the profit of the whole bank.
Inserting the market-based $FTP_t^M$ instead of $FTP_t^o$ in Equations 2 and 3, the optimal loan $(L_t^M)$ and deposit $(D_t^M)$ volumes are obtained when $FTP_t^M$ equals the bank’s $NMR(L_t)$ and its $MC(D_t)$, respectively. As long as $FTP_t^M$ is determined on an efficient financial market, the decisions related to the bank’s interest rates on customer loans $(r_{i,t}^c)$ and deposits $(r_{i,t}^d)$ may be delegated to its branch managers. Provided that treasury can settle all imbalances on the efficient financial market without incurring any additional costs, the bank’s overall profit would still be the sum of the profits of the branches, i.e. $\pi = \pi_L + \pi_D$. If treasury were taking liquidity positions and/or incurring additional costs for the bank, the profit of treasury ($\pi_T$) must be taken into account as well. The overall profit of the bank would then be: $\pi = \pi_L + \pi_D + \pi_T$.

3. FTP and the liquidity premium

If the bank considers its exposure to liquidity risk too high, treasury can indirectly influence decisions made at the branch level concerning optimal volumes and interest rates on loans as well as on deposits by adding-on a liquidity premium ($P_t > 0$) to the market-based $FTP_t^M$. To maintain efficient ALM, the slope of the swap curve must be continuously up-dated whenever the bank’s overall exposure to liquidity risks changes. This is particularly emphasized should the bank become a net provider of funds with a specific maturity $t$ on the financial market due to an adjustment of the $FTP_t^M$ with $P_t > FTP_t^o - FTP_t^M$.

Assuming that the optimal loan volume ($L_t^M$) for Branch $L$ exceeds the optimal deposit volume ($D_t^M$) for Branch $D$ when treasury applies the market-based $FTP_t^M$ on internal sales and purchases of funds with maturity $t$. This suggests that treasury must finance the difference ($L_t^M - D_t^M$) on the financial market. Hence, the bank is a net borrower of maturity $t$ funds and exposed to a liquidity risk motivating treasury to impose a liquidity premium $P_t$. Assume that $P_t$ reflects the estimated opportunity cost of the additional liquidity risk taken on by the bank when Branch $L$ accepts a new loan $l_t$ with maturity $t$ and/or the reduced risk for the bank when Branch $D$ attracts the corresponding deposit $d_j$. Clearly, $L_t^M$ and $D_t^M$ will no longer be optimal volumes for the branches when $P_t$ is added-on the market-based $FTP_t^M$. With the new internal price of funds ($FTP_t^o = FTP_t^M + P_t$), both Branch $L$ and Branch $D$ maximize their profits at new optimal loan ($L_t^o$) and deposit ($D_t^o$) volumes, respectively. These volumes are derived by replacing $FTP_t^o$ with $FTP_t^M$ in Equations 2 and 3. As $P_t > 0$, $L_t^o < L_t^M$ and $D_t^o > D_t^M$.

As long as $FTP_t^p < FTP_t^o$, the bank will continue to be a net borrower on the market of maturity $t$ funds and, thus, $L_t^p > D_t^p$. If $FTP_t^p = FTP_t^o$, the bank’s liquidity risk exposure would be fully eliminated ($L_t^p = D_t^p$). A problem arises if $P_t$ is set so that $FTP_t^p > FTP_t^o$. The bank would then become a net provider of maturity $t$ funds to the financial market, i.e. $L_t^p < D_t^p$. The problem for the bank is that treasury cannot sell its surplus of these funds ($D_t^p - L_t^p$) on the market at the used $FTP_t^p$. It has to accept the interest rate on the market on which $FTP_t^M$ is based. Hence, the liquidity premium paid to Branch $D$ on “surplus” deposits incurs a cost that must be borne by treasury and the bank as a whole. In the standard case, with linear internal demand and supply curves, the incurred cost for treasury equals the surplus of maturity $t$ funds times the liquidity premium paid out, i.e. $P_t(D_t^p - L_t^p)$. 


Given linear internal demand and supply curves, Equations 4-7 show how an imposed liquidity premium (\(P_t\)) affects, ceteris paribus, profits on the branch level, for treasury and for the bank as a whole:

\[
\Delta \pi_L = \int_{L_t^T}^{L_t^M} \tau(L_t) - \omega(L_t) - FTP_t^M dL_t - \pi_t \cdot L_t^P = \pi_t(-\frac{1}{2}(L_t^M - L_t^I) - L_t^I) = -\frac{1}{2}\pi_t(L_t^M + L_t^I), \quad (4)
\]

\[
\Delta \pi_D = \pi_t \cdot D_t^P + \int_{D_t^M}^{D_t^M} (FTP_t^M - \tau(D_t) + \omega(D_t))dD_t = \pi_t\left(D_t^P - \frac{1}{2}(D_t^M - D_t^H)\right) = \frac{1}{2}\pi_t(D_t^M + D_t^H), \quad (5)
\]

\[
\Delta \pi_T = (FTP_t^M - FTP_t^P)(D_t^P - L_t^I) = (FTP_t^M - (FTP_t^M + P_t))(D_t^P - L_t^I) = -\pi_t(D_t^P - L_t^I), \quad (6)
\]

\[
\Delta \pi = \Delta \pi_L + \Delta \pi_D + \Delta \pi_T = -\frac{1}{2}\pi_t(L_t^M + L_t^I) + \frac{1}{2}\pi_t(D_t^M + D_t^H) - \pi_t(D_t^P - L_t^I) \Rightarrow \quad (7)
\]
\[
\Delta \pi = -\frac{1}{2}\pi_t\left((D_t^P - L_t^I) + (L_t^M - D_t^H)\right)
\]

Clearly, the add-on of a liquidity premium will affect the profits of Branch \(L\) negatively and Branch \(D\) positively. Treasury will increase its profits if the bank remains a net borrower on the financial market of maturity \(t\) funds, whereas it will break-even if \(L_t^P = D_t^P\) and make a loss if the bank becomes a net funds provider on the market. For the bank as a whole, the liquidity premium will affect, ceteris paribus, its risk-unadjusted profit negatively. Provided that \(P_t\) accurately mirrors the opportunity cost of the bank’s exposure to liquidity risk, the effect on its risk-adjusted profit will be positive. However, the risk-adjusted profit is positively affected only as long the bank remains a net funds borrower on the financial market. If the bank becomes a net provider of funds with the specific maturity \(t\) targeted (i.e. \(L_t^P < D_t^P\)) it would imply “over-compensation” of its exposure to liquidity risk and that treasury has set a too high \(P_t\). Such a premium is unlikely to be sustainable over time.

In theory, a too high \(P_t\) must at least be lowered to: \(\hat{P}_t = \frac{\pi_t}{1 + \left(\frac{(D_t^P - L_t^I)}{(L_t^M - D_t^H)}\right)}\). This implies that \(FTP_t^P = FTP_t^P\), which makes it optimal for Branch \(L\) to purchase all available deposits from Branch \(D\). \(\text{10}\) Thereby, the change in the bank’s overall (risk un-adjusted) profit will be limited to the shaded area in the diagram to the right in Figure 1. Assuming no additional transaction costs, this area (which sums-up to: \(\Delta \pi = -\frac{1}{2}\pi_t\left(\frac{(L_t^M - D_t^H)^2}{(L_t^M - D_t^H)^2} + (L_t^M - D_t^H)^2\right) = -\frac{1}{2}\hat{\pi}_t(L_t^M - D_t^H)\)) reveals the bank’s theoretical cost to completely eliminate its exposure to liquidity risk caused by loans and deposits with maturity \(t\). \(\text{11}\)

On efficient financial markets, the liquidity premium \(\hat{P}_t\) materializes after a process of adjustments in the bank’s swap curve and (fixed- and floating-rate) borrowing costs. In reality, this process can take time. An alternative approach is to let treasury apply a so-called dual FTP to accomplish a reduction of an observed liquidity risk exposure. For instance, if treasury adds-on \(P_t\) to the market-based \(FTP_t^M\) only when selling funds with maturity \(t\) to Branch \(L\) and not when purchasing the corresponding deposits from Branch \(D\), the bank’s liquidity risk exposure will be reduced in accordance with the elasticity of the market demand for loans. While it will be optimal for Branch \(L\) to decrease its lending to \(L_t^P\), the deposit volume \(D_t^H\) will remain optimal for Branch \(D\). This suggests the following net (risk-unadjusted) profit change for the bank: \(\Delta \pi = -\frac{1}{2}\pi_t(L_t^M - L_t^P)\). If the market loan demand is more inelastic than the market supply for deposits, i.e. \(|\delta \text{MNR}(L_t)| > |\delta \text{MC}(D_t)|\), the negative effect on the bank’s risk-unadjusted profit would be lower if treasury instead does the opposite and
charges \( FTP^M_t \) on the maturity \( t \) funds sold to Branch \( L \) and adds-on a liquidity premium \( P_t \) to \( FTP^M_t \) when purchasing deposits from Branch \( D \). The net effect on the bank’s overall risk-unadjusted profit would then be: \( \Delta \pi = -\frac{1}{2} P_t (D_t^P - D_t^M). \)

Until now we have dealt with scenarios where the bank operates as a monopoly/monopsony or under competitive market conditions. However, as discussed in the introduction, this is seldom the case in practice. If the bank instead operates on a market characterized by oligopolistic and/or oligopsonistic competition, where the market demand curve for loans \( (M(L_t)) \) or the supply curve for deposits \( (M(D_t)) \) are “kinked”, treasury is likely to have to resort to one of the alternative approaches just mentioned. This will be further discussed and elaborated on in the next section.

4. FTP under oligopolistic and oligopsonistic competition

Under Cournot, Bertrand and Stackelberg competition, the framework for analyzing potential effects of a liquidity premium is essentially the same as in Section 3. However, the framework is another on oligopoly and oligopsony markets that are characterized by Sweezy-like competition. Then, all competing banks (branches) are facing kinked demand and/or supply curves on the retail markets of maturity \( t \) loans and deposits, respectively. The banks’ focus is set mainly on gaining and/or defending market shares. On one hand, either an increase in the lending rate \( (r^1_{t,t}) \) nor a decrease in deposit rate \( (r^2_{t,t}) \) of one bank is likely to be followed by the other competing banks. This is because these banks would then miss an opportunity to gain market share by keeping their, from a bank customer perspective, more attractive interest rates. On the other hand, the competing banks will be more urgent to respond actively whenever another bank decreases its lending rate or increases its deposit rate. If the banks remain passive, they would risk losing market share. This implies higher price rigidity on the retail bank market. However, as in other oligopolies and oligopsonies with Sweezy-like competition, a general change of the market interest rate that is the basis for the \( FTP^M_t \) used by treasury when purchasing and selling maturity \( t \) funds may very well alter the optimal volumes for all competing banks.

Oligopolistic competition and liquidity premium

Price stickiness of banks may have implications for treasury’s internal pricing of liquidity risk exposures. In the case of oligopolistic competition, the internal demand curve for loans with maturity \( t \)—i.e. the net marginal revenue \( (NMR(L_t)) \) of Branch \( L \)—breaks into two parts under the kink of the loan demand curve. This is shown in Equation 8:

\[
NMR(L_t) = \begin{cases} 
NMR^-(L_t) & \text{for } L_t < L^K_t, \\
NMR^+(L_t) & \text{for } L_t > L^K_t,
\end{cases}
\]  

(8)

where

- \( NMR^-(L_t) = \text{Net marginal revenue of loans with maturity } t \text{ for volumes below } L^K_t \)
- \( NMR^+(L_t) = \text{Net marginal revenue of loans with maturity } t \text{ for volumes above } L^K_t \)
- \( L^K_t = \text{Loan volume at the kink of the demand for loans with maturity } t \).
\( \text{NMR}(L_t) \) is not defined at the “kink” loan volume \( (L_t^K) \), because the market demand for loans below \( (M^-(L_t)) \) and above \( (M^+(L_t)) \) \( L_t^K \) exhibits different elasticities. Loan demand is steeper and, thus, more inelastic above than below this volume suggesting that \( \frac{\delta M^-(L_t)}{\delta L_t} < \frac{\delta M^+(L_t)}{\delta L_t} \).

Accordingly, the slopes of \( \text{NMR}^-(L_t) \) and \( \text{NMR}^+(L_t) \) will also differ resulting in a “gap” between the two parts. The greater the divergence of the loan demand elasticities (slopes), the greater the gap and the likelihood that it is optimal for \( \text{Branch L} \) to refrain from reducing its loan volume if treasury adds-on a liquidity premium \( P_t \) when settling imbalances of maturity \( t \) funds.

In a bank that uses a market-based \( \text{FTP}^M_t \), the optimal loan volume \( L_t^M \) equals \( L_t^K \) implying that \( \text{NMR}^-(L_t^K) < \text{FTP}^M_t < \text{NMR}^+(L_t^K) \). Assuming that the bank is a net borrower of maturity \( t \) funds on the financial markets (i.e. \( L_t^M > D_t^K \)) and that treasury imposes a liquidity premium \( P_t \) to reduce the bank’s exposure to liquidity risk. If \( P_t < \text{NMR}^-(L_t^K) - \text{FTP}^M_t \), the optimal volume for \( \text{Branch L} \) will be insensitive to the premium. This is because \( \text{FTP}^M_t < \text{FTP}^P_t < \text{NMR}^-(L_t^K) \) and, hence, \( L_t^K = L_t^M \). However, the add-on of \( P_t \) would make it optimal for \( \text{Branch D} \) to increase its deposit volume to \( D_t^P \). If then \( D_t^P > L_t^M \), the bank would become a net provider of maturity \( t \) funds on the market causing a net cost for the bank. Given a linear internal supply curve, the bank would suffer a profit reduction equal to: \( \Delta \pi = \frac{\pi}{2} P_t (D_t^P - D_t^K) \).

In this situation, treasury should instead adopt the alternative approach with dual \( \text{FTP} \) and add-on the less high optimal liquidity premium \( (\hat{P}_t) \) only when purchasing maturity \( t \) funds from \( \text{Branch D} \). Adding-on \( \hat{P}_t = P_t \left( 1 + \frac{(L_t^M - D_t^K)}{(D_t^P - D_t^K)} \right) \), the optimal deposit volume for \( \text{Branch D} \) equals the optimal loan volume for \( \text{Branch L} \), i.e. \( D_t^P = L_t^M \). This reduces the net cost for the bank to: \( \Delta \pi = -\frac{\pi}{2} \hat{P}_t (L_t^M - D_t^K) = \frac{\pi}{2} \hat{P}_t \left( \frac{(L_t^M - D_t^K)^2}{(D_t^P - D_t^K)} \right) \).

### Oligopsonistic competition and liquidity premium

In the case of oligopsonistic competition, where competing banks are facing a kinked supply curve, each bank will respond actively (passively) on any other bank’s increase (decrease) of the interest rate offered to deposit customers. In accordance with the prevalent strategy on an oligopoly market, no bank can gain market shares by adopting an aggressive pricing strategy. Hence, the banks will be reluctant to take any action unless other banks do. Due to the price stickiness on the deposit market, the internal supply curve \( (MC(D_t)) \) splits into two separate parts \( (MC^-(D_t)) \) and \( MC^+(D_t)) \) under the kink of the deposit supply curve \( (M(D_t)) \), where \( \frac{\delta MC^-(D_t)}{\delta D_t} > \frac{\delta MC^+(D_t)}{\delta D_t} \). This gives Equation 9:

\[
MC(D_t) = \begin{cases} 
MC^-(D_t) & \text{for } D_t < D_t^K, \\
MC^+(D_t) & \text{for } D_t > D_t^K,
\end{cases}
\]

where \( MC^-(D_t) = \text{Marginal cost of deposits with maturity } t \text{ for volumes below } D_t^K \), \( MC^+(D_t) = \text{Marginal cost of deposits with maturity } t \text{ for volumes above } D_t^K \), \( D_t^K = \text{Deposit volume at the kink of the supply for deposits with maturity } t \).

Analogously with the corresponding oligopoly case on the loan market, the resulting gap at the kink volume \( D_t^K \) could here make the optimal deposit volume \( D_t^P \) of \( \text{Branch D} \) more or
less insensitive to the imposition of a liquidity premium $\bar{P}_t$ on maturity $t$ funds. By definition $D^M_t$ equals $D^k_t$ at the market-based $FTP^M_t$ and, hence, $MC^-(D^k_t) = FTP^M_t < MC^+(D^k_t)$. This suggests that $D^M_t$ will remain the optimal deposit volume for Branch $D$ given any imposed $\bar{P}_t < MC^+(D^k_t) - FTP^M_t$. Such a liquidity premium paid by treasury will just generate additional revenue for Branch $D$. Provided that treasury also charges the additional $\bar{P}_t$ on maturity $t$ funds sold to Branch $L$ and the bank remains a net borrower of these funds on the financial market, i.e., $L^o_t \geq D^M_t$, treasury will not make any loss. However, its profit would be higher if it instead focuses on reducing the bank’s lending by adopting the alternative dual FTP approach and add-on $\bar{P}_t$ only when selling maturity $t$ funds. From the perspective of the bank, this approach is preferable if $FTP^o_t < MC^+(D^k_t)$. Depending on the slope of the internal demand curve ($NMR(L_t)$), the bank could then either remain a net borrower or become a net provider of maturity $t$ funds on the financial market when treasury adds-on $\bar{P}_t$ when selling the funds to Branch $L$. In the former case ($L^o_t > D^M_t$), the optimal liquidity premium ($\bar{P}_t$) would be lower than treasury’s $\bar{P}_t$, whereas it would be higher in the case $L^o_t < D^M_t$.16 Given a linearly decreasing internal demand curve, $\bar{P}_t = \bar{P}_t \left(1 + \frac{(L^o_t - n^{M_t})}{(L^o_t - L^M_t)}\right)$ in both cases. This gives the following net effect on the bank’s overall risk-unadjusted profit: $\Delta \pi = -\frac{1}{2} \bar{P}_t (L^M_t - D^M_t) = \frac{1}{2} \bar{P}_t \left(\frac{L^M_t - D^M_t}{L^o_t - L^M_t}\right)^2$.17

Simultaneous oligopolistic and oligopsonistic Sweezy-like competition

In the case of oligopolistic Sweezy-like competition on the retail loan market at the same time as there is the corresponding oligopsony on the retail deposit market, the benefit for the bank of having low-cost access to efficient financial markets is further emphasized. With no or costly access to the markets, like during the financial crisis, the settlement of imbalances between outstanding loans and deposits may cause the bank substantial losses in market shares. Market shares can also be foregone when treasury adds-on a liquidity premium to the market-based $FTP^M_t$.

The case of simultaneous Sweezy-like oligopolies and oligopsonies is distinguished from other market conditions in that optimal loans and deposit volumes may remain unchanged even if the bank were to face an increase in the current market-rate of maturity $t$ funds. On one hand, the degree of this robustness depends on the differences in the elasticities of the loan demand and deposit supply curves below and above their respective kink volumes ($L^k_t$ and $D^k_t$) and, thus, the implied gap between the internal demand curves $NMR^-(L_t)$ and $NMR^+(L_t)$ as well as between the internal supply curves $MC^-(L_t)$ and $MC^+(L_t)$. Under these conditions, a liquidity premium – whether dual or not – might have no effect on the bank’s optimal volumes of either loans or deposits. If the bank is lending-oriented at the current market-based $FTP^M_t$ (i.e. $L^k_t > D^k_t$), it will remain an equally large net borrower on the market of maturity $t$ funds unless treasury imposes a high enough liquidity premium $\bar{P}_t$. Either $\bar{P}_t > NMR^-(L^k_t) - FTP^M_t$ or $\bar{P}_t > MC^+(D^k_t) - FTP^M_t$, or both. If $\bar{P}_t < FTP^o_t - FTP^M_t$, the bank would still remain a net borrower on the financial market, though.

In the case of lending-oriented bank, $FTP^o_t$ must here by definition be greater than either $NMR^-(L^k_t)$ or $MC^+(D^k_t)$, or both. In the first case, where $MC^+(D^k_t) > FTP^o_t > NMR^-(L^k_t)$, an imposed $\bar{P}_t < MC^+(D^k_t) - FTP^M_t$ by treasury will not give any of the branches an incentive to change their current volumes implying that $L^o_t = L^M_t$ and $D^o_t = D^M_t$. Given an objective to fully eliminate the bank’s exposure to liquidity risk in maturity $t$ funds, treasury should adopt dual
FTP and add-on the liquidity premium \( \hat{\theta}_t = FTP_t^o + FTP_t^M \) only when purchasing funds from Branch D. In the second case, i.e. when \( NMR^- (L^F_t) > FTP_t^o > MC^+ (D_t) \), there will be no changes in volumes as long as treasury sets \( \mathcal{P}_t < NMR^- (L^F_t) - FTP_t^M \). Again, treasury should use dual FTP, but now the liquidity premium \( \hat{\theta}_t = FTP_t^o + FTP_t^M \) should only be added-on when selling maturity \( t \) funds to Branch L. Finally, in the third case, where either \( FTP_t^o > MC^+ (D^F_t) > NMR^- (L^F_t) \) or \( FTP_t^o > NMR^- (L^F_t) > MC^+ (D_t) \), treasury should abandon dual FTP and add-on the optimal liquidity premium \( \hat{\theta}_t = FTP_t^o - FTP_t^M \) both when purchasing and selling maturity \( t \) funds.

The third case suggests that the same \( FTP_t^o \) should be used as when there is no low-cost access for the bank to efficient financial markets. Thus, \( L^o_t = D^o_t = D^o_t \). After adjusting the loan volume to \( L^o_t \) and the deposit volume to \( D^o_t \), the bank’s loan demand curve and deposit supply curve will instead be kinked at \( L^o_t \) and \( D^o_t \), respectively. This results in parallel shifts of both its internal demand curve and its internal supply curve. The former shifts downwards above the kink on the loan demand curve, whereas the latter shifts upwards below the kink of the deposit supply curve. This implies that the bank’s loss of market shares on the retail loans and savings markets would not be regained if treasury thereafter lowered its liquidity premium or changed back to the market-based \( FTP_t^M \). This suggests that the add-on of \( \hat{\theta}_t \) will be costly for the bank both in the short and long run.

5. Concluding remarks

After the global financial crisis, practitioners, regulators and scholars advocate an adjustment of banks’ FTP that also covers the cost of liquidity risk exposures. Our analysis lends support to previous work (see e.g. Dermine, 2011; Grant, 2011; Cadamagnani et al., 2015), which suggest the addition of a term liquidity premium to the FTP as an extra charge on internal funds to finance lending, and a corresponding compensation when providing internal funds through deposits. We show that this strategy is optimal for the bank as long as it is (i) a net funds borrower on the financial market and (ii) is not operating under Sweezy-like competition. In the short-run, banks that fail to implement the liquidity premium underestimate the cost of funds and provide inappropriate incentives for the branches. In the long-run, such behaviour may also have systemic effects if the mispricing of funds results in retaliations from other banks (i.e. similarly to the post-crisis bank behaviour).

If the net funds providing bank is operating under Sweezy-like competition, our analysis demonstrates that it is optimal for the bank to differentiate the liquidity premium and, thus, use different internal prices on internal lending and borrowing. In particular, when operating on an oligopoly market where banks are highly price rigid, it is optimal for the bank that treasury adds-on a premium only when purchasing funds, whereas the opposite holds for the oligopoly case. Moreover, if price stickiness is prevalent on both the loans and deposits markets, the add-on of a liquidity premium may be too costly for the bank and not motivated at all.

Finally, our analysis also has implications for monetary policy and particularly the pass-through effect of policy interest rates. While there are several factors determining the extent to which policy rates are passed through to the economy (Burgstaller and Scharler, 2010), the FTP is bound to matter. If the banks overestimate their liquidity premiums, economic activity may be hampered. Similarly, if the banks underestimate these premiums, it may result in
excessive lending. In addition, if banks are operating under Sweezy-like competition, our analysis indicates that monetary policy actions taken by the Central bank may be less effective since the kinks will inhibit the pass-through of policy rate changes to the real economy.

References


---

1 E.g., Kimball, 1997; Norris, 2002; Kipkalov, 2004; Kawano, 2005; Levey, 2008; Hanselman, 2009; Button, Pezzini, Rossiter, 2010; Grant, 2011; Cadamagnani, Harmohan, Tangri 2015.

2 In 2011-2015 Bank of England published a series of papers on the topic of FTP, explaining how it is currently being altered within the banks to meet the new requirements and to better mimic risk. However, more formal analyses of the implications
of new FTP regulations are still missing and this paper aims to fill that gap (see Cadamagnani et al. (2015) for a review of this series).

1 Both deposit rates and operational costs are adjusted to compensate for that all deposits made is not available for lending.

2 One of the very reasons for decentralization is to delegate decision-right to the organizational unit which possesses the most relevant information to make the best decision for the business. We refer to the unrealistic case of complete information just to illustrate the optimal FTP\textsubscript{L}.

3 Where the former is decreasing, and the latter is increasing, with volume, i.e. \( \frac{\delta \text{NMR}(L_D)}{\delta t} < 0 \) and \( \frac{\delta \text{MC}(D_t)}{\delta t} > 0 \).

4 The separation theorem does not hold for all decisions concerning loans and deposits. Some decisions affect both loans and deposits. See Dermine (2011) for an extended discussion on how to handle dependent decisions.

5 As the internal demand and supply curves are linear, the increase of the bank’s overall profit (\( \Delta \pi \)) of adopting a market-based FTP is: \( \Delta \pi = \frac{\delta \text{FTP}_M - \delta \text{FTP}_D}{\delta t} (L_D^M - L_D^D) \).

6 Neither NMR\textsubscript{L,D} nor MC\textsubscript{D} need to be linear. However, it is required that \( \frac{\delta \text{NMR}(L_D)}{\delta t} < 0 \) and \( \frac{\delta \text{MC}(D_t)}{\delta t} > 0 \) when approaching the market-based FTP\textsubscript{L}.

7 In our analysis, we assume that treasury will always act in the interest of the bank.

8 This means that \( L_t = D_t = \frac{(a - a_{\text{M}})}{\text{M}_t} \).

9 This implies a cost reduction by: \( \gamma \cdot \text{NMR}\textsubscript{L,D} \left( (D_t^M - L_t^M) + (L_t^D - D_t^D) \right) - \gamma \cdot \text{MC}\textsubscript{D} \left( \frac{(a - a_{\text{M}})}{\text{M}_t} \right) \).

10 If the liquidity exposure of the bank were to be fully eliminated with the alternative dual FTP approach, treasury must add on its liquidity premium to fully eliminate the bank’s exposure to liquidity risk in maturity \( t \) funds to Branch \( L \) and Branch \( D \), respectively. Adding on \( \text{FTP}_D = \text{FTP}_D^\ast \left( \frac{(a - a_{\text{M}})}{\text{M}_t} \right) \) would make it optimal for Branch \( L \) to adjust its loan volume in accordance with \( L_D^M \). Alternatively, adding-on \( \text{FTP}_D = \text{FTP}_D^\ast \left( \frac{(a - a_{\text{M}})}{\text{M}_t} \right) \) would make it optimal for Branch \( D \) to adjust its deposit volume in accordance with \( L_D^M \). As both \( \gamma > \psi \) and \( \gamma > \psi^\ast \), it would be less costly for the bank to fully eliminate its exposure to liquidity risk if treasury instead adds-on \( \psi \), both when selling and purchasing these funds internally.

11 “Deposits have become much more important today, but deposits are like gasoline: you can increase the price and if you are successful others will follow. Accordingly, we have raised our FTPs on deposits and our external prices to attract more deposits but in fact the business units do not seem to experience an increased inflow and we will most likely go back to our previous prices – the price-elasticity is just not that high on deposits.” Top Manager at a large Swedish bank.

12 Where AR\textsubscript{L,D} = \( \int_t L_t^D \, dL_t / L_t \) and AC\textsubscript{L,D} = \( \int_t D_t^D \, dD_t / D_t \).

13 If the bank remains a net borrower on the market of maturity \( t \) funds given \( \text{P}_t \), i.e. \( D_t^D < L_t^D \), treasury must instead increase its liquidity premium to fully eliminate the bank’s exposure to liquidity risk in maturity \( t \) funds. If \( \text{FTP}_D^\ast < \text{NMR}\textsubscript{L,D} \left( L_t^D \right) \), the same formulas are used to determine \( \psi \) and the net profit effect. Hence, \( \psi = \text{FTP}_D^\ast \left( \frac{(a - a_{\text{M}})}{\text{M}_t} \right) \) and \( \Delta \pi = -\gamma \cdot \text{FTP}_D^\ast \left( L_t^M - D_t^M \right) \).

14 Should \( \text{FTP}_D^\ast > \text{NMR}\textsubscript{L,D} \left( L_t^D \right) \), treasury should not use dual FTP but instead add-on the optimal liquidity premium \( \psi = \text{FTP}_D^\ast - \text{FTP}_D^\ast \) both when purchasing and selling of maturity \( t \) funds internally.

15 If \( \text{FTP}_D^\ast > \text{MC}\textsubscript{D} \left( D_t^D \right) \), the bank should not use dual FTP but use the optimal liquidity premium \( \psi = \text{FTP}_D^\ast - \text{FTP}_D^\ast \) on both internal sales and purchases of maturity \( t \) funds.