

THE UNIVERSITY OF
SYDNEY

Economics Working Paper Series

2020 – 12

A Macroeconomic Theory of Banking Oligopoly

Mei Dong, Stella Huangfu and Hongfei Sun

October 2020

A Macroeconomic Theory of Banking Oligopoly*

Mei Dong[†]
University of Melbourne

Stella Huangfu[‡]
University of Sydney

Hongfei Sun[§]
Queen's University

Chenggang Zhou
Canada Mortgage and Housing Corporation[¶]

October 9, 2020

Abstract

We study the behavior and macroeconomic impact of oligopolistic banks in a tractable environment with micro-foundations for money and banking. Our model features oligopolistic banks, which resembles the structure of the banking sector observed in most advanced economies. Banks interact strategically where they compete against each other in terms of the volume of loans to make. We find that it is welfare-maximizing to have the banking sector as oligopolistic, *i.e.*, to have a small number of large banks. In addition, moderate inflation improves welfare because it helps to ease congestion in the banking sector.

JEL Categories: D83, E43, E44, E51, E52, G21, G32, L26

Keywords: banking, oligopoly, liquidity, market frictions

*We are grateful for feedback from George Alessandria, Yan Bai, Francesco Carli, Allen Head, Charlie Kahn, Thor Koepl, Jim McGee, Stephen Williamson, Randall Wright, and seminar participants at Australian National University, Bank of Canada, Bank of Japan, Keio University, Queen's University, University of Queensland, University of Rochester, 2017 Canadian Economic Association Annual Meeting, 2017 Society for Economic Dynamics Meeting, 2017 Summer Workshop on Money, Banking, Payment and Finance, 2017 Workshop of Australasian Macroeconomics Society, and 2018 Centre for Banking and Financial Stability Workshop at Deakin University. Dong acknowledges financial support under Australian Research Council's DECRA scheme (project number: DE120102589). Sun acknowledges support from Social Sciences and Humanities Research Council (SSHRC).

[†]Email address: mei.dong@unimelb.edu.au. Mailing Address: Department of Economics, Faculty of Business and Economics, Carlton, Victoria, Australia, 3053.

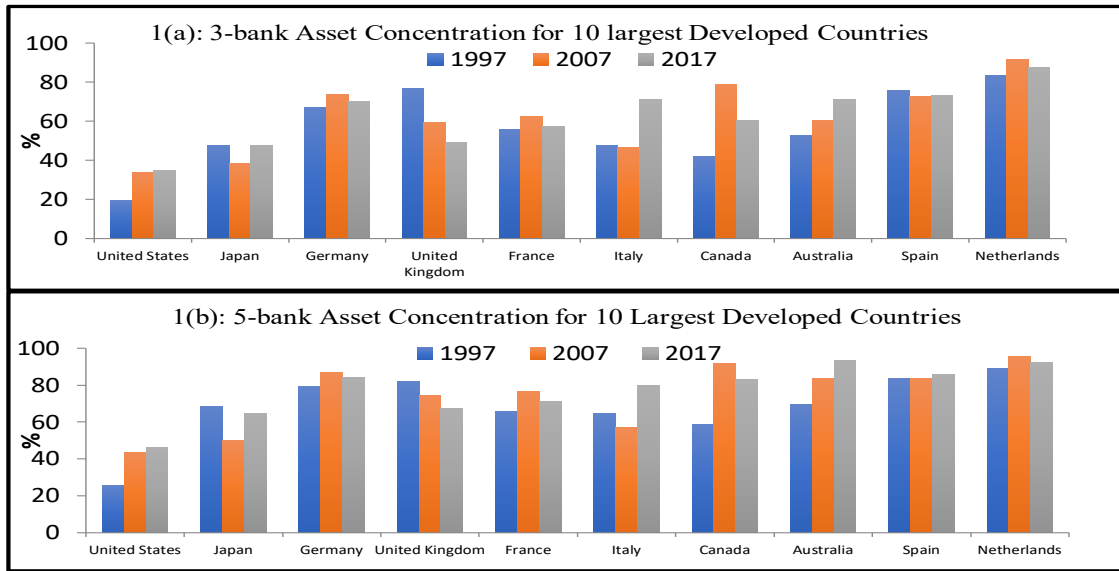
[‡]Email address: stella.huangfu@sydney.edu.au. Mailing Address: School of Economics, University of Sydney, NSW, Australia, 2006.

[§]Email address: hfsun@econ.queensu.ca. Mailing Address: Department of Economics, Queen's University, Dunning Hall, Room 309, 94 University Avenue, Kingston, ON, Canada, K7L 3N6.

[¶]Email address: czhou@cmhc-schl.gc.ca. Mailing Address: Canada Mortgage and Housing Corporation, 700 Montreal Rd, Ottawa, Ontario, Canada, K1A 0P7.

1 Introduction

We study the behavior and the macroeconomic impact of oligopolistic banks in a tractable environment with micro-foundations for money and banking. Nowadays in many countries, the banking sector is clearly an oligopoly, in the sense that it consists of a few large banks who control a significant proportion of the banking business across the country. This is demonstrated by Figures 1(a) and 1(b), which respectively report the concentration of the banking sector measured by the fraction of total bank assets controlled by the top 3- or top 5-largest banks in ten largest advanced economies. For instance, the top five banks in the U.S. accounted for about 46 per cent of total bank assets in 2017.



Data source: World Bank Global Financial Development Database

Figure 1: Concentration of the Banking Sector

Since the 2008 Financial Crisis, the importance of banking sector has prompted new research interests in macroeconomic models of banking. However, existing literature on banking often ignore the fact that the banking sector consists of a few large banks and treats the banking sector as being composed of either one monopoly bank or a continuum of competitive banks with no market power. As is well known from the Cournot competition model, oligopolistic firms could behave differently from monopolistic or competitive firms, which would further lead to different economic outcomes. To overcome this gap between theory and evidence, it is necessary to model banks as oligopolists and the market structure of the banking sector (i.e., the number of banks) as endogenously

determined.

In this paper, we construct a dynamic general equilibrium model to study how the structure of the banking sector affects the macroeconomy and welfare. We model the banking sector as oligopolistic where individual banks are large in the sense that each bank has a large number of loan officers working to provide loans. Banks compete for shares in the credit market by choosing the amount of loan officers at work, which is essentially choosing the volume of new loans to create every period. On one hand, each bank's incentive to issue loans becomes stronger as more banks compete against each other. On the other hand, the amount of deposits available per bank shrinks with more banks at play. The two forces together create tension in banks' liquidity needs and affect banks' profitability.

Holding the number of banks exogenous, we find that welfare responds to the number of banks non-monotonically in a hump shape. The highest welfare corresponds to a small, yet greater than one, number of banks. In other words, *it is optimal to have an oligopolistic banking sector*. The intuition is the following: as the number of banks rises, more intense oligopolistic competition induces more loans being issued in aggregate. As a result, more firms get funding and more workers find jobs, all of which contribute to a positive effect on welfare. However, having more banks is also associated with two negative welfare effects: First, it becomes increasingly more difficult for banks to make loans as more banks crowd in the credit market. This is reflected in a declining probability of loan-issuing by each officer of a bank. Second, banks become liquidity constrained because each bank receives a smaller share of aggregate deposits with more banks in competition. This worsens banks' financial situation. Overall, competition in the banking sector improves welfare only to a certain extent. This is a unique result out of studying banking oligopoly.

We then endogenize the number of banks to examine the relationship among inflation, bank competition and welfare. Our main finding is that moderate inflation improves welfare. This is because there is over-entry of banks in equilibrium. Inflation tends to reduce banks' profits, which deters bank entry and thus helps ease congestion in the banking sector. In our baseline case, welfare is maximized when annual inflation is 1.53%. Nevertheless, with more severe inflation the standard welfare cost of inflation dominates

the positive effect associated with easing bank entry, and thus welfare declines with inflation.

Our theoretical framework is partly built upon Wasmer and Weil (2004) (henceforth WW) and Berentsen, Menzies and Wright (2011) (henceforth BMW). In particular, WW has search frictions in the credit and labor markets, while BMW features frictional labor and goods markets (without a credit market). Although WW focuses on how credit market conditions affect labor market outcomes, there is an important piece missing in the model in that the source of the amount to lend by creditors is unaccounted for. In contrast, we specifically fill in this missing piece of the puzzle by modeling a full-fledged banking sector who gather funds from workers before lending to firms. Moreover, this banking sector is the main focus of our paper as we aim to study how decisions of oligopolistic banks influence the aggregate economy.

Our way of modeling the banking sector is novel, in the sense that (i) we model banks as oligopolists in a tractable macro environment with solid micro-foundation for money and banking, and (ii) we address the funding-liquidity problem faced by oligopolist banks. The advantage of our modelling choice is that the frictional environment allows us to examine the consequence of liquidity problems faced by oligopolist banks. Credit market frictions naturally give rise to well-defined intensive margin (*i.e.*, the size of a loan) and extensive margin (*i.e.*, the volume of loans issued by a bank) of credit. Goods market frictions, as in Lagos and Wright (2005), provide a rigorous micro-foundation for a medium of exchange, *i.e.*, money. In addition, such a structure is tractable as it renders the equilibrium distribution of money degenerate. This is particularly convenient for us in that we aim to examine the macroeconomic consequences of banking. Moreover, labor *and* goods market frictions together help us build a solid micro-foundation for banking. Because of labor market frictions, it may take several periods before a firm recruits a worker and starts generating revenue. This uncertainty of loan repayment prevents workers from lending to firms directly because workers need money to consume in the frictional goods market. In contrast, banks can avoid this uncertainty and intermediate between a large set of workers and firms even though workers and firms face idiosyncratic matching risks. Thus, banks can meet withdrawal demand from workers who need money for consumption.

With these features, we bring our unique contribution to the vast literature on banking. The theoretical banking literature mainly consists of microeconomic models of banking. However, macroeconomic models of banking have attracted much interest in recent years especially since the 2008 Financial Crisis.¹ Compared to the standard macroeconomic banking models, our paper focuses on the long-run effect of having an oligopolistic banking sector. It is closest in relation to the theories of money, credit and banking with explicit micro-foundations for money and banking, *and* focuses particularly on bank lending activities. In this regard, the literature narrows down to a few papers such as Berentsen et al. (2007) and Sun (2007, 2011).² Among these papers, only Sun (2007) considers banks as oligopolists. In Sun (2007), banks serve as delegated monitors and there arises a micro-foundation for money due to lack of double coincidence of wants. In the model, the number of banks is exogenous. Competition among banks strictly improves welfare by stimulating aggregate lending. In contrast, our model allows for an endogenous number of banks. Moreover, we examine the consequence of liquidity problem faced by oligopolist banks and find that oligopolistic competition among banks does not always improve welfare.

More recently, Corbae and D’Erasmus (2013) endogenizes the bank size distribution using a quantitative model of banking. They model banking competition as a Stackleberg game and focus on the dynamics of bank size distribution during business cycles. Huang (2019) develops a model of banking where an optimal number of banks arises due to trade-off between monitoring costs and rewards to banks. The decentralized allocation features excess entry of banks, which can be corrected by appropriately-designed banking policies. While we argue that regulating the number of banks can improve welfare, our model features a frictional goods market that allows us to examine how inflation interacts with the structure of the banking sector and welfare. Head et al. (2020) propose a model with imperfect competition in the consumer-loan market based on the noisy-search model as in Burdett and Judd (1983). Their model generates endogenous concentration of banking

¹See Gorton and Winton (2003) and Freixas and Rochet (2008) for surveys of microeconomics models of banking. Important contributions to macroeconomic models of banking include Gertler and Kiyotaki (2010, 2015) and Gertler et al. (2016). These papers study potential banking crisis in the spirit of Diamond and Dybvig (1983).

²Other papers with micro-foundations of money and banking include Cavalcanti and Wallace (1999), Williamson (2004), He *et al.* (2005, 2008), Williamson and Wright (2010a, 2010b), Gu *et al.* (2013). Nevertheless, these models do not have oligopolistic banks.

sector and heterogenous loan rates among consumers. In contrast, we model imperfect competition among banks as Cournot competition.

The remainder of this paper is organized as follows. Section 2 describes the economic environment. We provide value functions and derive the optimal decisions made by workers, firms and banks in Section 3. Section 4 characterizes the banking equilibrium. Section 5 explores the quantitative implications of our model. The final section offers some concluding remarks.

2 Model Environment

Time is discrete and has infinite horizons. Each time period t consists of *three* subperiods. As shown in Figure 2, a centralized goods market (GM) and a *credit* market (CM) are simultaneously active in the first subperiod. A frictional labor market (LM) is active in the second subperiod. A decentralized goods market (DM) is active in the third subperiod. All goods are perishable across periods. All agents discount each period at rate $\beta \in (0, 1)$. Discounting occurs following DM activity and preceding GM/CM activity.

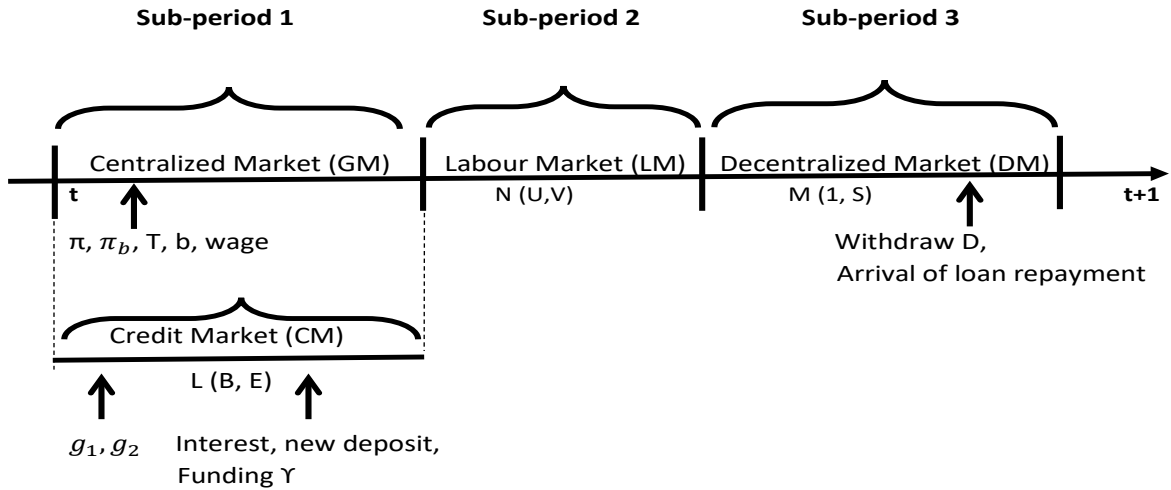


Figure 2: Timing of Events

The economy is populated by three types of agents: workers, firms and banks, which are indexed by h , f and b , respectively. The measure of infinitely lived workers is one. Workers supply labor and consume in both GM and DM. They also own all the firms

in the economy. The measure of firms is sufficiently large, although not all firms will be active at any point in time. Firms produce output by hiring labor while also paying wages and dividends to workers. Firms require funding from banks before they can hire workers and start production. There is also a banking sector composed of N banks, where $N \geq 1$ is finite. Banks take in demand deposits from workers and lend to firms.

Due to anonymity and the lack of double coincidence of wants, agents require a means of payment in DM. An intrinsically worthless, perfectly divisible and storable object, fiat money, serves this purpose. The quantity of money at the end of subperiod 1 of period t is M_t . We assume that $M_t = (1 + \pi)M_{t-1}$, where π is the growth rate of money supply. New money is injected by lump-sum transfers to workers at the beginning of subperiod 1 in GM. The price of goods in terms of money in GM is p_t . We restrict our attention to stationary equilibria in which the real value of aggregate money balances M_t/p_t is constant. It implies that $p_t = (1 + \pi)p_{t-1}$ and π is the inflation rate. To facilitate exposition, from now on we adopt the following notation: M denotes money balances in period t , and \hat{M} represents money balances in period $t+1$, similarly for other observables. Throughout the paper, we use GM goods as the numeraire. We denote the respective values of entering GM, CM, LM and DM as W^j , I^j , U^j , and V^j , where $j \in \{h, f, b\}$ is the agent type.

Subperiod 1: In subperiod 1 of every period t , a frictionless GM and a frictional CM are simultaneously active. Each bank in CM has a large number of loan officers who assist in lending. At the beginning of this subperiod, each bank selects from a pool of idle loan officers. The selected loan officers receive instructions on how to negotiate with matched firms over loan terms, before being sent into the CM. New firms can enter CM at a sweat cost $k > 0$. Loan officers and new firms are matched randomly and bilaterally. Matching is governed by the function $\mathcal{L}(B, E)$, where $B = \sum_{n=1}^N B_n$ is the total measure of loan officers from all banks, B_n is the measure of loan officers from bank n , and E is the measure of new firms in CM.

A new firm faces a cost of $\gamma > 0$ units of GM goods every time it posts vacancies in LM. As such, the loan contract requires the bank to lend γ to the firm in every subperiod 1 until the firm hires a worker. From that point on, the firm will start repaying a real

amount a every period until the job terminates, which occurs with probability s every period. The job separation shock occurs at the onset of each subperiod 2.

Subperiod 2: In subperiod 2, firms (with funding from banks) and unemployed workers are randomly and bilaterally matched in LM. Upon meeting, the two parties bargain over the real wage, and production takes place immediately afterwards. Matching is governed by the function $\mathcal{N}(u, v)$, where u is the measure of unemployed workers and v is the measure of vacancy. The output in a match is denoted by y .

Subperiod 3: In subperiod 3, all workers and firms with output y are randomly matched in the DM. Again, matching is bilateral and is governed by the function $\mathcal{M}(1, 1 - u)$, where the measure of workers is 1 and the measure of firms with output y is $1 - u$. Once matched, the parties engage in Kalai bargaining over the terms of trade (q, d) , where a real money balance d is used to purchase an amount q of DM goods.³ The DM goods can be paid with only fiat money. Banks are not allowed to issue inside money, but workers can withdraw their deposits to obtain fiat money. We assume that all matching functions have constant returns to scale and satisfy the usual assumptions.

3 Values and Decisions

In this section, we describe the value functions for each type of agents and solve the bargaining problems in each market. The value functions for workers and firms are standard in models with frictional labor and goods market. The decisions faced by banks are unique to our framework and thus we provide a detailed description of the banking problem.

3.1 Workers

Consider a worker h with employment status $e \in \{0, 1\}$ who enters GM with real deposit balances z , where $e = 1$ indicates that the worker is currently employed, and $e = 0$ indicates that the worker is unemployed. The worker maximizes her utility by choosing the amounts of GM consumption, X , and new demand deposit to be made in the current

³We choose Kalai bargaining according to Kalai (1977) because it helps to simplify the theoretical results without loss of generality.

subperiod, \hat{z} . Note that workers will choose asset portfolios consisting of only demand deposits as long as the nominal interest rate on demand deposits $i_d > 0$. The worker's value of entering GM is given by

$$\begin{aligned} W_e^h(z) &= \max_{X, \hat{z}} \{X + (1 - e)\ell + U_e^h(\hat{z})\} \\ \text{st. } X + \hat{z} &= ew + (1 - e)b + \Pi + \rho\Pi_{b,-1} - T + (1 + i_d)z, \end{aligned} \quad (1)$$

where ℓ is the utility of leisure, w is the real wage, b is unemployment insurance (UI), Π and $\Pi_{b,-1}$ are respectively dividends from firms and banks, and T is the lump-sum transfer due to taxation and inflationary purposes. Note that dividends from banks ($\Pi_{b,-1}$) are paid to workers as money balances and thus are subject to inflation with $\rho = 1/(1 + \pi)$. As is standard with the Lagos-Wright type of environment, the quasi-linear preference renders the model very tractable. Notice that W_e^h is linear in z : $dW_e^h/dz = 1 + i_d$. Moreover, from (1), the optimal choice of \hat{z} is independent of z . As will be shown later, the portfolio of the worker is independent of her employment status e . As a result, every worker has the same \hat{z} when exiting GM.

In LM, the value function for a worker is constructed according to her employment status,

$$U_1^h(z) = sV_0^h(z) + (1 - s)V_1^h(z), \quad (2)$$

$$U_0^h(z) = \lambda^h V_1^h(z) + (1 - \lambda^h)V_0^h(z), \quad (3)$$

where s is the job separation rate, and $\lambda^h = \mathcal{N}(u, v)/u = \mathcal{N}(1, v/u)$ is the probability for an unemployed worker to find a job. For simplicity, we assume that if a match is destroyed, a worker cannot gain employment until the next LM.

In DM, the worker meets a firm with probability α^h , where $\alpha^h = \mathcal{M}(1, 1 - u)$. If they trade, the worker pays a real balance d for q units of goods and thus the value function is given by

$$V_e^h(z) = \alpha^h [u(q) + \beta \hat{W}_e^h(\rho z - \rho d)] + (1 - \alpha^h) \beta \hat{W}_e^h(\rho z). \quad (4)$$

Note that real money balances that are carried into the next period are inflation-adjusted with ρ . The function $u(q)$ is the utility that the worker obtains by consuming q units of DM goods. We assume that $u(0) = 0$, $u' > 0$ and $u'' < 0$.

For a worker, the above value functions of the three respective subperiods can be collapsed into one Bellman equation of the entire time period. After repeated substitution, we obtain the following recursive problem:

$$W_e^h(z) = I_e + (1 + i_d)z + \beta \mathbb{E} \hat{W}_{\hat{e}}^h(0, 0) + \max_{\hat{z}} \{ \hat{\alpha}^h [u(\hat{q}) - \beta(1 + i_d)\rho \hat{d}] - \hat{z}[1 - \beta(1 + i_d)\rho] \}, \quad (5)$$

where the expectation is with respect to next period's employment status \hat{e} , $I_1 = w + \Pi + \rho\Pi_{b,-1} - T$ and $I_0 = b + \ell + \Pi + \rho\Pi_{b,-1} - T$. Note that the optimal choice of \hat{z} does not depend on the employment status of a worker. An interior choice of \hat{z} satisfies

$$\hat{\alpha}^h \left[u'(\hat{q}) \frac{\partial \hat{q}}{\partial \hat{z}} - \beta \rho (1 + i_d) \frac{\partial \hat{d}}{\partial \hat{z}} \right] = 1 - \beta \rho (1 + i_d). \quad (6)$$

3.2 Firms

There are three types of firms: firms not associated with any bank or worker (type 0), firms with a bank loan but not a worker (type 1) and firms who have both a bank loan and a worker (type 2). We first consider the value of a type-0 firm at the beginning of subperiod 1, represented by I_0^f . A type-0 firm can enter CM at a sweat cost $k > 0$ to search for a loan officer for funding. The firm will be matched with a loan officer with probability $\phi^f = \mathcal{L}(B, E)/E = \mathcal{L}(B/E, 1)$, and enter LM as a type-1 firm. If not matched, the firm remains type 0. Thus,

$$I_0^f = \max\{0, \phi^f U_1^f(a) + (1 - \phi^f)\beta \hat{I}_0^f - k\}. \quad (7)$$

Consider a type-1 firm with a loan contract of periodic repayment a (for short, contract a). With probability $\lambda^f = \mathcal{N}(u, v)/v = \mathcal{N}(u/v, 1)$, the firm finds a worker and immediately produces output y . The firm then enters DM as a type-2 firm. Otherwise, it remains type 1. Thus,

$$U_1^f(a) = \lambda^f V_2^f(a) + (1 - \lambda^f)\beta \hat{U}_1^f(a). \quad (8)$$

Only type-2 firms with output y enter DM to trade. In a match, the firm sells q units of DM goods, which costs the firm $c(q)$ in terms of GM goods and we have $c'(q) > 0$

and $c''(q) \geq 0$. The residual $y - c(q)$ is transformed into GM goods and sold in the subsequent GM. The cost function here captures the variable costs that firms incur in production. The corresponding value function is given by

$$V_2^f(a) = \alpha^f \beta \hat{W}_2^f(y - c(q), \rho d, a) + (1 - \alpha^f) \beta \hat{W}_2^f(y, 0, a), \quad (9)$$

where $\alpha^f = \mathcal{M}(1/(1-u), 1)$ is the probability of a firm meeting and trading with a worker in DM and $\hat{W}_2^f(\cdot, \cdot, \cdot)$ is the value of a type-2 firm at the beginning of subperiod 1 of $t + 1$.

In GM, a type-2 firm clears inventory x , makes loan repayment a , and pays wage w . The profit $x + z - w - a$ aggregating across all type-2 firms constitutes dividends Π paid to workers. Thus a type-2 firm's value function is given by

$$W_2^f(x, z, a) = x + z - w - a + s\beta \hat{I}_0^f + (1 - s) V_2^f(a). \quad (10)$$

For simplicity, we assume that the lending relationship ends once the employment relationship ends due to the separation shock. With probability s , the firm becomes type 0 again. Otherwise, the firm remains type 2.

3.3 Banks

Banks issue demand deposits to workers in each subperiod 1. The demand deposit is a *contract* between a bank and a worker, defined as follows.

Definition 1 *The demand deposit contract states that: (i) for a real balance z deposited in subperiod 1 of t , the worker is free to withdraw any amount less than or equal to z by the end of t ; (ii) The bank is to pay a net nominal interest rate, $i_d > 0$, on the worker's remaining real bank balance at the beginning of subperiod 1 of $t + 1$.*

Consider any bank $n \in \{1, 2, \dots, N\}$ in CM. The bank sends a measure B_n of loan officers to CM. Sending each loan officer incurs a real cost $\kappa > 0$. Let $\mathcal{A} = \{a : a \in [0, \infty)\}$ be the set of all possible contracts and $a \in \mathcal{A}$ represent a contract with a particular periodic repayment level, a . Moreover, let A represent the contract term negotiated between loan officers and firms in the *current* CM. Now define $g_{ni}(a)$ as the density of the distribution of all type- i contracts owned by bank n at the beginning of subperiod

1 of t , with $i = 1, 2$. In other words, $g_{ni}(a)$ represents the measure of type- i contracts owned by bank n that require a periodic repayment of a . Moreover, type-1 contracts refer to those that are in the *funding* stage (*i.e.*, firms searching in the LM) and the type-2 are those in the *repayment* stage (*i.e.*, firms making repayments to the bank).

Given that demand deposits are costly, the bank may have the incentive to retain some of its earnings to help alleviate the cost of funds in the upcoming periods. We call the retained earnings “bank savings”. Let K_n denote the real balance of bank savings at the end of subperiod 3 of $t - 1$, and \hat{Z}_n be the amount of demand deposits accepted by bank n in subperiod 1 of period t . Moreover, $D_n \in [0, \hat{Z}_n]$ is the expected amount of withdrawal in DM coming up in subperiod 3 of t . Finally, let \bar{B}_{-n} be the total measure of loan officers sent by all banks except n . That is, $\bar{B}_{-n} = \sum_{n \neq k=1}^N B_k$. The probability for a loan officer to meet a firm in CM is $\phi^b(B_n, \bar{B}_{-n}, E) = \mathcal{L}(B, E)/B = \mathcal{L}(1, E/B)$, where $B = B_n + \bar{B}_{-n}$. The periodic profit of bank n is given by

$$\begin{aligned} \Pi_{bn} = & \rho K_n + \int_{a \in \mathcal{A}} ag_{n2}(a) da + (1 - \kappa_d) \hat{Z}_n - \kappa B_n - (1 + i_d) \rho Z_{nD} \\ & - \gamma \left[\int_{a \in \mathcal{A}} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right] - D_n - \hat{K}_n - \kappa_f, \end{aligned} \quad (11)$$

where $\kappa_d \in (0, 1)$ is a variable cost associated with handling demand deposits and $\kappa_f > 0$ is a fixed cost. Note that κ_d is a sweat cost of the bank and does not affect the two liquidity constraints specified next. Moreover, $\int_{a \in \mathcal{A}} g_{ni}(a) da$ is the stock of type- i contracts held by bank n , and $B_n \phi^b(B_n, \bar{B}_{-n}, E)$ is the measure of new type-1 contracts created in the current subperiod 1. The bank’s profit Π_{bn} is to be rebated to workers as dividends at the beginning of subperiod 1 of $t + 1$.

Between the current subperiod 1 and the next, the bank has three sources of funds: its own savings, ρK_n , repayments from all type-2 contracts, $\int_{a \in \mathcal{A}} ag_{n2}(a) da$, and newly accepted demand deposits, \hat{Z}_n . Note that the bank’s own saving K_n is subject to the inflation tax because this is the balance at the end of subperiod 3 of $t - 1$ and can only take the form of fiat money.

Banking costs include the cost of handling demand deposits, $\kappa_d \hat{Z}_n$, the cost of sending loan officers to CM, κB_n , the cost of funding new and existing type-1 contracts, $\gamma \left[\int_{a \in \mathcal{A}} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right]$, the cost of paying principal and interests on

previously-issued demand deposits, $(1 + i_d) \rho Z_{nD}$, the cost of meeting demand of withdrawals, D_n , and finally, the fixed cost, κ_f . In addition, the bank will choose a new amount of own savings \hat{K}_n to put aside by the end of subperiod 3 of t .

We now proceed to describe the bank's decision problem. First, when making decisions, bank n takes A , the current loan term, as given because it is to be determined through negotiation between its loan officers and fund-seeking firms. When sending loan officers to CM, bank n informs its loan officers that the marginal value of creating a new loan to the bank is $\mathcal{V}_{n1}(a)$ given any a . The officers will carry out negotiations according to this instruction. Moreover, the bank also takes the competitive deposit rate, i_d , withdrawal demand in real terms, D_n , measure of new firms in CM, E , and all other banks' decisions as given. Second, there are four state variables for bank n 's decision problem. They are the stocks of type- i contracts, g_{ni} with $i = \{1, 2\}$, the total balance of demand deposits at the beginning of subperiod 1, Z_{nD} , and the savings from subperiod 3 of $t - 1$, K_n . Finally, the bank makes decisions on the measure of loan officers to send to CM, B_n , the amount of new demand deposits to take in, \hat{Z}_n , and the real balance to save by the end of subperiod 3, \hat{K}_n .

We assume that all loan repayments submitted by firms in subperiod 1 of t will arrive at the corresponding banks in subperiod 3 of t . This assumption is used to augment a potential *liquidity problem* for banks, resulting from mismatched timing of payments. It is common in reality that banks need to manage the flow of funds by keeping track of the timing of payments. We interpret the late arrival of repayments as a result of some exogenous reasons such as technical details, time difference, etc., but not because of any wrong-doing on the firms' part. In other words, we abstract from any incentive problem of firms. Banks have full knowledge of the arrival time of repayments, and make decisions accordingly.

Given $(A, i_d, D_n, E, \bar{B}_{-n})$, the bank chooses $(B_n, \hat{Z}_n, \hat{K}_n)$ to maximize the present value of current and future profits:

$$\mathcal{V}_n(g_{n1}, g_{n2}, Z_{nD}, K_n) = \max_{(B_n, \hat{Z}_n, \hat{K}_n)} \left\{ \Pi_{bn} + \beta \hat{\mathcal{V}}_n(\hat{g}_{n1}, \hat{g}_{n2}, \hat{Z}_{nD}, \hat{K}_n) \right\}$$

$$\text{st. } \rho K_n + \hat{Z}_n \geq \gamma \left[\int_{a \in A} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right] + \kappa B_n + (1 + i_d) \rho Z_{nD} + \kappa_f \quad (12)$$

$$\rho K_n + \hat{Z}_n + \int_{a \in A} a g_{n2}(a) da \geq D_n + \gamma \left[\int_{a \in A} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right] + \kappa B_n + (1 + i_d) \rho Z_{nD} + \kappa_f + \hat{K}_n \quad (13)$$

$$\hat{K}_n \geq 0 \quad (14)$$

$$\hat{Z}_{nD} = \hat{Z}_n - D_n \quad (15)$$

$$\hat{g}_{n1}(a) = \begin{cases} (1 - \lambda^f) [g_{n1}(a) + B_n \phi^b(B_n, \bar{B}_{-n}, E)], & \text{if } a = A \\ (1 - \lambda^f) g_{n1}(a), & \text{if } a \neq A \end{cases} \quad (16)$$

$$\hat{g}_{n2}(a) = \begin{cases} (1 - s) g_{n2}(a) + \lambda^f [g_{n1}(a) + B_n \phi^b(B_n, \bar{B}_{-n}, E)], & \text{if } a = A \\ (1 - s) g_{n2}(a) + \lambda^f g_{n1}(a), & \text{if } a \neq A. \end{cases} \quad (17)$$

Condition (12) is the liquidity constraint faced by bank n by the end of subperiod 1 of t . We refer to it as the *lending liquidity constraint* (lending constraint hereafter). The right-hand side of (12) consists of the costs and payments that must be covered by the bank within subperiod 1. They are lending costs associated with γ and κ , the payment of principal and interests on previous demand deposits, $(1 + i_d) \rho Z_{nD}$, and the fixed cost. The left-hand side of (12) consists of the available funds for covering the above costs and payments. There are two sources of such funds: bank savings at the beginning of subperiod 1, ρK_n and newly issued demand deposits, \hat{Z}_n . Because of the delay in arrival of loan repayments, demand deposits can be particularly helpful for relieving the liquidity condition of banks, since bank savings can be costly due to inflation.

Condition (13) is another liquidity constraint faced by bank n by the end of subperiod 3 of t , when workers withdraw their demand deposits to purchase DM goods. We refer to

this constraint as the *withdrawal liquidity constraint* (withdrawal constraint hereafter). Comparing with condition (12), the additional funds that the bank receives is the loan repayment $\int_{a \in \mathcal{A}} ag_{n2}(a) da$ that arrives in subperiod 3. On the expense side, there are two additional expenses faced by bank n : withdrawal demand D_n and bank's savings \hat{K}_n .

Condition (14) states that the bank's savings must be nonnegative. Condition (15) states that the total balance of demand deposits at the end of subperiod 3 of t is equal to the amount of deposits issued in subperiod 1 of t , \hat{Z}_n , minus the total amount of withdrawal in subperiod 3 of t , D_n . Conditions (16)-(17) are the laws of motion for the distributions of contracts owned by a bank. The distributions are over a , the repayment requirement of the contracts. In general, the stocks of contracts owned by a bank may have different loan requirements, depending on the economic conditions at the time of creating contracts. Recall that the bargained contracts have the term $a = A$ in the current CM. Therefore, at the beginning of period $t + 1$, the type-1 contracts with particular requirement $a = A$ consist of the previous and the newly-added type-1 contracts of $a = A$ that are not matched with a worker in period t . Similarly, the evolution of type-2 contracts with $a = A$ must include those existing type-2 contracts that are not hit by a separation shock, and those newly added to the pool of type-2 contracts. Finally, measures of the contracts with any requirement $a \neq A$ evolve in their due courses given LM shocks, without adding to existing stocks.

Let ε and ξ be the Lagrangian multipliers for the bank's liquidity constraints (12) and (13), respectively. Using the envelop conditions, we derive the optimal conditions for interior solutions of $(B_n, \hat{Z}_n, \hat{K}_n)$ as follows

$$\left[\phi^b(B_n, \bar{B}_{-n}, E) + B_n \frac{\partial \phi^b}{\partial B_n} \right] \frac{\frac{\beta \lambda^f (1+\xi) A}{1-\beta(1-s)} - \gamma(1+\varepsilon+\xi)}{1-\beta(1-\lambda^f)} = \kappa(1+\varepsilon+\xi), \quad (18)$$

$$\beta \rho(1+i_d) = 1 - \frac{\kappa_d}{1+\varepsilon+\xi}, \quad (19)$$

$$\beta \rho(1+\varepsilon+\xi) - (1+\xi) = 0. \quad (20)$$

It is possible that $\beta \rho(1+\varepsilon+\xi) - (1+\xi) < 0$, which implies $\hat{K}_n = 0$.

3.4 Bargaining Solutions

Let η_b be the loan officer's bargaining power in CM. Before sending them to CM, bank n instructs its loan officers the marginal value of a type-1 contract with a given term a , *i.e.*, $\mathcal{V}_{n1}(a) = \partial \mathcal{V}_n / \partial g_{n1}(a)$. In CM, loan officers and firms engage in Nash bargaining. If they agree on a deal, the loan officer brings the value $\mathcal{V}_{n1}(a)$ back to the bank, and the firm receives $U_1^f(a)$. If no deal, then the bank receives nothing and the firm retains the value $\beta \hat{I}_0^f$. Therefore, the bargaining solution of a splits the surplus between the two parties according to

$$\frac{\mathcal{V}_{n1}(a)}{U_1^f(a) - \beta \hat{I}_0^f} = \frac{\eta_b}{1 - \eta_b}. \quad (21)$$

The solution is solved from the following optimal condition:

$$\frac{\beta \lambda^f a(1 + \xi) - \gamma(1 + \varepsilon + \xi)[1 - \beta(1 - s)]}{\beta \lambda^f \{y + \alpha^f[\rho d - c(q)] - w - a\}} = \frac{\eta_b}{1 - \eta_b}. \quad (22)$$

In LM, firms and workers engage in Nash bargaining, taking the loan contract that the firm has with a bank as given. Effectively given a , the wage rate $w(a)$ solves

$$\frac{V_2^f(a) - \beta \hat{U}_1^f(a)}{V_1^h(z) - V_0^h(z)} = \frac{\eta_f}{1 - \eta_f}, \quad (23)$$

where η_f is the firm's bargaining power. If both parties agree, the firm receives $V_2^f(a)$ and the worker receives $V_1^h(z)$. Otherwise, the firm and worker continue with values $\beta \hat{U}_1^f(a)$ and $V_0^h(z)$, respectively. Due to the linearity of $W_e^h(z)$, the optimal condition is given by

$$\frac{(1 - \beta)[1 - \beta(1 - s - \lambda^h)]}{[1 - \beta(1 - s)][1 - (1 - \lambda^f)\beta]} \frac{y + \alpha^f[\rho d - c(q)] - w - a}{w - (b + \ell)} = \frac{\eta_f}{1 - \eta_f}.$$

In DM, firms and workers bargain over the terms of trade (q, d) according to the Kalai protocol. Let μ be the bargaining power of a worker. The surplus for a worker is $u(q) - \beta(1 + i_d)\rho d$. Similarly, the firm's surplus is given by $\beta[\rho d - c(q)]$. The Kalai bargaining solves

$$\max_{(q,d)} u(q) - \beta(1 + i_d)\rho d \quad \text{st.} \quad \frac{u(q) - \beta(1 + i_d)\rho d}{\beta[\rho d - c(q)]} = \frac{\mu}{1 - \mu} \quad \text{and} \quad d \leq z. \quad (24)$$

The resulting terms of trade depend on z if and only if the constraint $d \leq z$ binds.

Lemma 1 *The optimal bargaining solution in DM is*

$$d(z) = \begin{cases} z^*, & \text{if } z \geq z^* \\ z, & \text{if } z < z^* \end{cases}$$

$$q(z) = \begin{cases} q^*, & \text{if } z \geq z^* \\ g^{-1}(\beta\rho z), & \text{if } z < z^* \end{cases}$$

where $g(\cdot)$ is defined in (25) and q^* is implicitly defined by $u'(q^*)/c'(q^*) = \beta(1 + i_d)$ and $\beta\rho z^* = g(q^*)$,

$$\beta\rho d = g(q) \equiv \frac{(1 - \mu)u(q) + \mu\beta c(q)}{(1 - \mu)(1 + i_d) + \mu}. \quad (25)$$

The proof of the above lemma, as well as all other proofs, are provided in Appendix A. The results in this lemma are standard for variations of the Lagos-Wright environment. As will be shown later, in any equilibrium with money and banking, a worker will spend all of her real balances in a DM trade. That is, $d = z$ and $q = g^{-1}(\beta\rho z)$ because there is no precautionary motive of saving through holding fiat money as a result of quasi-linear preferences.

4 Equilibrium

4.1 Definitions

Let the distributions of type- i contracts aggregated across all N banks be denoted by $g_i(a) = \sum_{n=1}^N g_{ni}(a)$ for $i = 1, 2$. We have the following equilibrium definition:

Definition 2 *A symmetric banking equilibrium $(\hat{Z}_n, B_n > 0)$ consists of values*

$$\left\{ (\mathcal{V}_n)_{n=1}^N, (U_e^h, V_e^h, W_e^h)_{e=0,1}, (I_0^f, U_1^f, V_2^f, W_2^f) \right\},$$

decision rules $\left\{ (X, \hat{z}), (B_n, \hat{Z}_n, \hat{K}_n, D_n)_{n=1}^N \right\}$, measures $\{B, E, u, v\}$, prices and terms of trade $\{w(a), q(a), d(a), i_d\}$, and distributions $\{g_i(a)\}_{i=1,2}$ such that given policy (b, T, π) , the following are satisfied:

1. All decisions are optimal;
2. All bargaining results are optimal in (21), (23) and (24);

3. *Free entry of firms and banks is such that*

$$I_0^f = 0. \quad (26)$$

$$\Pi_{bn} = 0, \quad \forall n. \quad (27)$$

4. *Matching in the credit and labor markets satisfies:*

$$\mathcal{L}(B, E) = \mathcal{N}(u, v). \quad (28)$$

$$\mathcal{N}(u, v) = (1 - u)s. \quad (29)$$

5. *All competitive markets clear. In particular, clearing of money and demand-deposit markets requires:*

$$\frac{M}{P} = \sum_{n=1}^N \left\{ \int_{a \in \mathcal{A}} a g_{n2}(a) da + \rho K_n + \hat{Z}_n - \gamma \left[\int_{a \in \mathcal{A}} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right] - \kappa B_n - (1 + i_d) \rho Z_{nD} - \kappa_f \right\}. \quad (30)$$

$$\hat{z} = \sum_{n=1}^N \hat{Z}_n. \quad (31)$$

6. *Consistency: The aggregate distributions satisfy:*

$$\hat{g}_i(a) = \sum_{n=1}^N \hat{g}_{ni}(a), \quad i = 1, 2 \text{ and } \forall a,$$

where $\hat{g}_{ni}(a)$ obeys laws of motion (16)-(17);

7. *Symmetry: $(B_n, \hat{Z}_n, \hat{K}_n)$ are respectively the same for all n in any time period.*

Thus,

$$B = NB_n, \text{ and } \alpha^h \hat{z} = ND_n.$$

8. *Government balances budget:*

$$bu = T + \pi \frac{M}{P}. \quad (32)$$

Conditions 1-3 and 6-7 are self-explanatory. Condition 4 ensures that the inflows and outflows to the stock of type-1 firms (those with a bank contract but without a worker), and those to the stock of jobs are balanced in equilibrium. The money market clearing condition (30) is particularly worth mentioning. The right-hand side of (30) is the money

demand at the end of subperiod 1 of t . In particular, loan repayments by productive firms must be in the form of fiat money for two reasons: first, it is the only way for the repayments to arrive at the banks in subperiod 3 since real goods do not last across subperiods. Second, banks need cash to accommodate withdrawal demand in DM in subperiod 3. For the same reasons, any balance of a bank at the end of subperiod 1, $\rho K_n + \hat{Z}_n - \gamma \left[\int_{a \in \mathcal{A}} g_{n1}(a) da + B_n \phi^b(B_n, \bar{B}_{-n}, E) \right] - \kappa B_n - (1 + i_d) \rho Z_{nD} - \kappa_f$, must also be carried forward as cash. Given that both money and deposit markets clear, as in (30) and (31), the GM must also clear by Walras's law. Condition 8 stipulates that the UI program is supported by taxes imposed on workers and seigniorage.

Definition 3 *A stationary, symmetric banking equilibrium is one in which all real values and distributions are time-invariant. In particular, the contract term remains the same over time at $a = A$, and*

$$g_1(A) = \left(\frac{1}{\lambda^f} - 1 \right) \mathcal{L}(B, E) \quad (33)$$

$$g_2(A) = \frac{\mathcal{L}(B, E)}{s} \quad (34)$$

$$g_i(a) = 0 \quad \forall a \neq A \in \mathcal{A}.$$

Equations (33)-(34) are derived from (16)-(17) given $\hat{g}_{ni}(A) = g_{ni}(A)$ for all n and i in the steady state. Note that in the stationary equilibrium, all wages and terms of trade in DM are constant over time because of the constant contract term. That is, there is no dispersion of contract terms, wages or goods prices in the steady state. Let $w = w(A)$ represent the steady-state wage from this point on.

4.2 Characterizing the Stationary Banking Equilibrium

We focus on the case of $1 + \pi > \beta$, although we do consider the limit as $1 + \pi \rightarrow \beta$, which is the Friedman rule. We study the stationary equilibrium and therefore drop all time subscripts hereafter. First, we characterize workers' optimal choice of z as follows.

Proposition 1 *In the steady state, worker's optimal decisions are such that:*

- (i) *If i_d satisfies $\beta \rho (1 + i_d) = 1$, then $d < z$.*

(ii) If i_d satisfies $\beta\rho(1+i_d) < 1$, then $d = z = g(q)/\beta\rho$ and q satisfies

$$\alpha^h \frac{u'(q)}{g'(q)} + (1 - \alpha^h)(1 + i_d) = \frac{1}{\beta\rho}. \quad (35)$$

The intuition of Proposition 1 is straightforward. When $\beta\rho(1+i_d) = 1$, bank deposits generate real returns that compensate exactly for discounting. Thus, workers are indifferent between consuming the GM goods and making deposits. When the deposit rate more than compensates for discounting, workers will only deposit the necessary amount (for purchasing the DM goods) z into the bank.

Next, firm's entry condition (7) implies that in the steady state

$$\beta\phi^f \lambda^f \frac{y + \alpha^f [\rho d - c(q)] - w - A}{1 - \beta(1-s)} = k[1 - \beta(1 - \lambda^f)]. \quad (36)$$

Lastly, the free entry of banks implies $\Pi_{bn} = 0$ and endogenously determines the equilibrium number of banks N . Thus, N must satisfy

$$(\rho-1)K_n + [(1 - \rho(1+i_d))(1-\alpha^h) - \kappa_d] \hat{Z}_n + Ag_{n2}(A) - \gamma[g_{n1}(A) + B_n \phi^b(B, E)] - \kappa B_n - \kappa_f = 0. \quad (37)$$

We are now ready to characterize the banking equilibrium. For banking equilibria to exist, we assume that $k < \{\beta[y + g(q^*)/\beta - c(q^*) - (b + \ell)]/[1 - \beta(1-s)] - \gamma\}/[(1 - \eta_f)(1 - \beta)/\eta_f + 1/(1 - \eta_b)]$ and $\kappa_f < [1 - \beta(1-s)]\{k\eta_b/(1 - \eta_b) + \gamma\}/\beta - [g(q^*)(1 - \kappa_d)/\rho(1 - \beta)]$ because otherwise, it would be too costly for firms or banks to enter the CM according to firm's entry condition (7) or bank's zero-profit condition (37).

Recall that ε and ξ are the Lagrangian multipliers for the bank's liquidity constraints (12) and (13), respectively. The following proposition establishes possible types of banking equilibria when the number of banks is exogenous.

Proposition 2 *If a banking equilibrium exists, a stationary banking equilibrium must have $d = z$ and the equilibrium deposit rate i_d satisfies (19). Moreover, six possible types of banking equilibrium can occur depending on the number of banks N and other parameter values.*

(I) $\varepsilon = \xi = K_n = 0$. Neither of the two liquidity constraints binds for banks. Banks choose $K_n = 0$. This type of steady state can be determined by solving (B, E, u, v, A)

from (18), (22), (28), (29), and (36);

(II) $\varepsilon > 0$ and $\xi = K_n = 0$. Lending constraint (12) binds but withdrawal constraint (13) does not bind. Banks choose $K_n = 0$. This type of steady state can be determined by solving $(B, E, u, v, A, \varepsilon)$ from (12) with equality, (18), (22), (28), (29), and (36);

(III) $\varepsilon > 0$, $\xi = 0$ and $K_n > 0$. Lending constraint (12) binds but withdrawal constraint (13) does not bind. Banks carry a strictly positive balance of K_n . This type of steady state can be determined by solving $(B, E, u, v, A, \varepsilon, K_n)$ from (12) with equality, (18), (22), (20) (28), (29), and (36);

(IV) $\varepsilon > 0$, $\xi > 0$ and $K_n > 0$. Both constraints (12) and (13) bind for banks. Banks carry a strictly positive balance of K_n . This type of steady state can be determined by solving $(B, E, u, v, A, \varepsilon, \xi, K_n)$ from (12) with equality, (13) with equality, (18), (20), (22), (28), (29), and (36);

(V) $\varepsilon = 0$, $\xi > 0$ and $K_n = 0$. Lending constraint (12) does not bind but withdrawal constraint (13) binds. Banks choose $K_n = 0$. This type of steady state can be determined by solving (B, E, u, v, A, ξ) from (13) with equality, (18), (22), (28), (29), and (36);

(VI) $\varepsilon > 0$, $\xi > 0$ and $K_n = 0$. Both constraints (12) and (13) bind for banks. Banks choose $K_n = 0$. This type of steady state can be determined by solving $(B, E, u, v, A, \varepsilon, \xi)$ from (12) with equality, (13) with equality, (18), (22), (28), (29), and (36).

The presence of the constraints naturally generates different types of banking equilibria. Figure 3 shows how equilibrium switches types as N increases. In type I equilibrium, neither constraint binds. Banks have sufficient funding to cover loan issuance in the first subperiod and deposit withdrawal in the third subperiod. This is more likely to occur when the number of banks is low so that each bank has a relative large share of the market and is not liquidity constrained. The proof of Proposition 2 shows that a non-binding lending constraint (12) implies $K_n = 0$. Intuitively, a non-binding lending constraint implies that banks have enough liquidity and so there is no need to rely on costly internal savings.

As the number of banks rises, each bank receives a smaller share of deposits, which tightens banks' lending constraint and withdrawal constraint. Either a type II equilibrium in which banks' lending constraint (12) binds or a type V equilibrium in which banks'

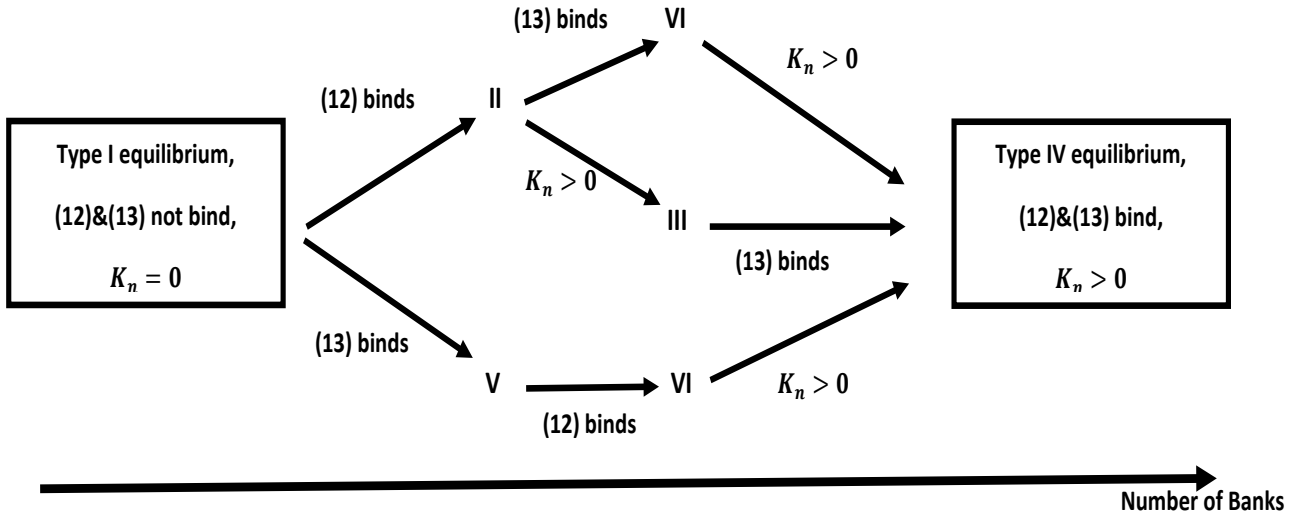


Figure 3: Types of Equilibria

withdrawal constraint (13) binds will arise. In either case, banks do not yet resort to their own savings as a source of funding ($K_n = 0$). The intuition is this: with more banks, it is true that increased demand for deposits from banks will stimulate the aggregate deposits supplied by all workers. But banks will be more likely to face binding constraints with a higher number of banks in competition, as long as the drop in deposit market share of each bank dominates the induced increase in deposits supply.

When the lending constraint (12) binds first, the economy is in a type II equilibrium. In this case, as the number of banks increases, the tension of liquidity needs associated with having more banks eventually pushes banks to either have a binding withdrawal constraint (13) or to save on their own. With two binding constraints and no own saving, the economy will be in a type VI equilibrium. Banks are constrained but do not find it optimal to save on their own. However, it is also possible that banks begin to save ($K_n > 0$) before the withdrawal constraint (13) binds. In this case, the economy switches from a type II to a type III equilibrium.

When the withdrawal constraint (13) binds first, the economy is in a type V equilibrium. In this case, having more banks in the economy can push banks to have a binding lending constraint. In other words, the economy switches from a type V to a type VI equilibrium. Note that banks do not save as long as the lending constraint (12) does not

bind.

Eventually, when the number of banks is extremely high, competition is so severe that both constraints bind and banks must rely on their own savings to fund future lending ($K_n > 0$). As a result, the economy switches into a type IV equilibrium from either a type VI or a type III equilibrium. We will conduct quantitative analysis in the next section to explore the nature of these different types of banking equilibria. Our numerical results suggest that types V and VI equilibria do not exist under reasonable parameters.

The characterization of banking equilibria with an endogenous N is straightforward: each banking equilibrium is solved as described in Proposition 2 with the addition of (37) to solve for N . With an endogenous N , there are three possible types of banking equilibria corresponding to types I – III above. Types IV–VI equilibria cannot occur because the zero-profit condition implies that (13) never binds, i.e., $\xi = 0$.

5 Quantitative Analysis

We now use quantitative analysis to further explore the properties of banking equilibria and examine the effects of inflation. To conduct this exercise, we calibrate the model using U.S. data in 2003.⁴

5.1 Parameterization

The model period is set to be one quarter. Following the macro-labor literature, the numerical analysis in this section uses the following specifications. In LM, the matching function is assumed to be $\mathcal{N}(u, v) = uv/(u^\phi + v^\phi)^{\frac{1}{\phi}}$. In DM, the utility function has the functional form of $u(q) = Aq^{1-\alpha}/(1-\alpha)$, and the cost function is $c(q) = q^\theta$, where $\theta \geq 1$. The matching function in DM is assumed to be $\mathcal{M}(1, 1-u) = (1-u)/(2-u)$. Thus, the matching probabilities of buyers (workers) and sellers (firms) are as follows: $\alpha^h = (1-u)/(2-u)$, and $\alpha^f = 1/(2-u)$. Moreover, the matching function in CM is assumed to be $\mathcal{L}(B, E) = BE/(B^\psi + E^\psi)^{\frac{1}{\psi}}$.

Given these specifications, the parameters to be determined include the preference parameters (β, A, α) , the matching function parameters (ϕ, ψ) , the technology parameters

⁴We choose year 2003 for two reasons. First, to exclude the periods of Global Financial Crisis (GFC) and post-crisis recovery due to impact of GFC on the economy. Second, some key statistics on banking sector are only available from a 2003 financial survey.

$(y, s, \eta_f, \eta_b, \mu, \theta)$, cost parameters $(\gamma, \kappa_d, \kappa_f, \kappa, k)$, the policy parameters (b, ℓ, ρ) , and the number of banks operating in the economy (N) .

Preference parameters. The discount factor β is set to 0.992 to match the annual real interest rate in the U.S., which is 3 percent.

Matching function parameters. We set the LM matching function elasticity parameter ϕ to 1.25, which is the same value as in Petrosky-Nadeau et al. (2018). Following Petrosky-Nadeau and Wasmer (2015), the CM matching function elasticity ψ is set to 1.35.

Technology parameters. The productivity in a match y is set to 2. The job separation rate is $s = 0.05$, which is used in Andolfatto (1996) and is within the range of the estimates reported by Davis et al. (2006). Next, the firm's bargaining power in LM η_f is set to 0.95, which is used by Hagedorn and Manovskii (2008). We set the loan officer's bargaining power in CM to $\eta_b = 0.62$, following Petrosky-Nadeau and Wasmer (2015). Worker's share of trade surplus in DM μ is set to 0.13, which is used in Bethune et al. (2015). Finally, the curvature parameter θ in the cost function $c(q)$ is set to 1.35.

Cost parameters. According to Silva and Toledo (2009), the average labor cost of hiring one worker is 3 to 4.5 percent of quarterly wages of a new hire. Hence, we set $\gamma = 0.045 * w$. Due to the lack of direct evidence, bank's variable cost of handling demand deposits is set to $\kappa_d = 0.001$. We will conduct a robustness check by varying its value to 0.005.

Policy parameters. The annual CPI inflation rate in 2003 reported by the Federal Reserve Bank is 2%, which results in quarterly value of $\rho = 0.995$. The calibrations of the two remaining policy parameters, i.e., the value of leisure ℓ and UI benefits b , are more controversial in the search literature. Shimer (2005) sets $\ell = 0$ and finds that the unemployment rate is not responsive enough to changes in productivity. Alternatively, Hagedorn and Manovskii (2008) argue that if this parameter is calibrated to match hiring costs, $b + \ell$ should be set to 0.95, resulting in a response of unemployment to productivity that is more consistent with the data. Since the effects of inflation on unemployment in our model are channeled through changes in firm revenue, we follow Hagedorn and Manovskii (2008) and set $b = 0.5$ and $\ell = 0.45$.

The number of banks N plays a crucial role in our model. The empirical counterpart

to N can be viewed as the inverse of the Herfindahl-Hirschman Index (HHI) for the U.S. banking sector. The HHI is a commonly accepted measure of market concentration, which is calculated by taking the sum of squared market share of each firm competing in a specific market. The inverse of the HHI yields the number of “effective competitors,” or the number of equal-sized firms that would produce an equivalent HHI score.⁵ Black and Kowalik (2017) find that the HHI of U.S. banks for the deposit market is 0.12 in 2003, which suggests that $N = 1/0.12 \simeq 8$. This is considered as a moderately concentrated market by the U.S. Department of Justice (DOJ). As robustness checks, we consider two alternative values of N , i.e., $N = 4$ (highly concentrated) and $N = 12$ (not concentrated).⁶

Notation	Parameter	Value
β	Discount factor	0.992
ϕ	Elasticity of the LM matching function	1.25
ψ	Elasticity of the CM matching function	1.35
y	Productivity	2
s	Separation rate	0.05
η_f	Firm’s bargaining power in the LM	0.95
η_b	Loan officer’s bargaining power in the CM	0.62
μ	Worker’s share of trade surplus in the DM	0.13
θ	Curvature parameter in $c(q)$	1.35
γ	Vacancies posting costs	$0.045w$
κ_d	Bank’s variable cost of handling demand deposits	0.001
b	UI benefits	0.5
ℓ	Value of leisure	0.45
ρ	Inverse of inflation rate	0.995
N	Number of banks	8

Table 1 summarizes the parameter values used in the calibration. The remaining parameters ($A, \alpha, k, \kappa_f, \kappa$) are calibrated using the following strategy: the two preference

⁵For banking sector, the HHI can be calculated by the sum of squared market share by each bank in a given market. For example, suppose there are ten banks in a market. The five largest banks each has 15 percent market share, and the remaining 5 banks each has 5 percent market share. The resulting HHI would be $5(0.15)^2 + 5(0.05)^2 = 0.125$. Under this definition, an extremely competitive market would have an HHI approaching zero, while a pure monopoly would have an HHI of 1. One complication arising in the case of banking sector is related to the definition of market share. In the literature, the market share has been approximated by a bank’s asset share, deposit share, or loan share. It would be ideal to use the HHI computed based on loan share in our calibration. However, we use the HHI based on deposit share as proxy due to data availability.

⁶The DOJ divides the spectrum of market concentration into three roughly delineated categories that can be broadly characterized as unconcentrated (HHI below 0.1), moderately concentrated (HHI 0.1 – 0.18), and highly concentrated (HHI above 0.18).

parameters (A, α) are chosen to match (i) retail markup of 40% reported by Faig and Jerez (2005), (ii) money demand $M/(pY) = 0.4$ in 2003. In the model,

$$\text{markup} = \frac{\rho d}{c'(q)q} - 1, \text{ and}$$

$$\frac{M}{pY} = \frac{[1 - \rho(1 + i_d)(1 - \alpha^h)]z + \rho N K_n - \gamma[g_1(A) + B\phi^b(B, E)] - \kappa B - N\kappa_f + Ag_2(A)}{(1 - u)y + \alpha^h[\rho d - c(q)]}.$$

The three cost parameters (k, κ_f, κ) are calibrated to match the following targets: (i) the average unemployment rate is 0.06 in 2003, (ii) the 2003 National Survey of Small Business Finances reports that 68% of small business have access to lines of credit provided by banks. We therefore set $\phi^f = 0.68$, and (iii) the financial sector's share of aggregate value added used by Petrosky-Nadeau and Wasmer (2013) is 2.5%. In our model, this share is

$$\Sigma = \frac{N(\Pi_{bn} + \kappa_d \hat{Z}_n)}{(1 - u)y + \alpha^h[\rho d - c(q)]}.$$

Notations	Description	Baseline	Highly Concentrated	Not Concentrated	Cost
A	Level parameter in $u(q)$	1.328	1.328	1.328	1.418
α	Curvature parameter in $u(q)$	0.321	0.321	0.321	0.275
k	Firm's cost of entering CM	6.397	6.397	6.397	6.397
κ_f	Bank's fixed cost ($\ast 10^{-2}$)	1.505	4.349	0.855	1.551
κ	Cost of sending loan officers to CM	7.286	6.701	7.480	7.286

Table 2 reports the calibrated values of the parameters $(A, \alpha, k, \kappa_f, \kappa)$. The first column corresponds to the baseline calibration described above. For robustness checks, we also present three alternative calibrations in the other columns. In the first alternative, labeled Highly Concentrated, we set the number of banks N to 4 which represents a more concentrated banking sector. In the second alternative, labelled Not Concentrated, N is set to 12. In the last column, labelled Cost, we set the bank's variable cost of handling demand deposits κ_d to 0.005. It should be noted although these alternatives are arbitrarily set, they are useful for illustrating how the results depend on the parameters.

5.2 Results

Using the calibrated parameters, we compute the equilibrium of the model. We summarize our numerical results into two categories: one set of results are obtained taking the number of banks N as given. For the second set of results, we endogenize N by including the bank's free-entry condition.

5.2.1 Equilibrium with an Exogenous N

We first analyze the equilibrium in which the number of banks is taken as given. This is helpful to understand how the number of banks can affect the economy. It will also help better understand the equilibrium with an endogenous N , whose results will be presented next.

Figure 4 illustrates how a set of equilibrium variables vary with the number of banks in the baseline model.⁷ As indicated in Proposition 2, in theory, six types of equilibria may exist. However, using calibrated parameter values, we find that types V and VI equilibria do not occur. In fact, as the value of N increases, our model features four non-overlapping regions in which types I to IV equilibria exist. In Figure 4, the four different types of banking equilibria are represented by four different colors. The blue color represents type I equilibrium ($1.020 \leq N \leq 10.979$) in which neither the lending constraint nor the withdrawal constraint binds and banks do not save. As N increases, the lending constraint tightens and the banking equilibrium switches to type II. The red color represents type II equilibrium ($10.980 \leq N \leq 11.356$) in which only the lending constraint binds and banks do not save. Any further increase in N induces banks to save on their own and the banking equilibrium switches to type III. The green color represents type III equilibria ($11.357 \leq N \leq 11.468$) in which only the lending constraint binds and banks start to save. Lastly, the banking equilibrium switches to type IV as the withdrawal constraint binds. The black color represents type IV equilibrium ($11.469 \leq N \leq 12.628$) in which both constraints bind and banks save.⁸

Figure 5 depicts how welfare varies with the number of banks in the economy for the

⁷For simplicity, we treat the number of banks as a continuous variable when we compute these results.

⁸Types II, III and IV equilibria exist for relatively smaller ranges of N . Figure 4 provides an overview of the curves for $N \geq 1$ and Figure B1 in the Appendix provides a partial view of equilibrium variables depicted in Figure 4. The partial view is meant to provide a better view of the segment where the equilibrium switches types as N changes ($10.95 \leq N \leq 12.628$).

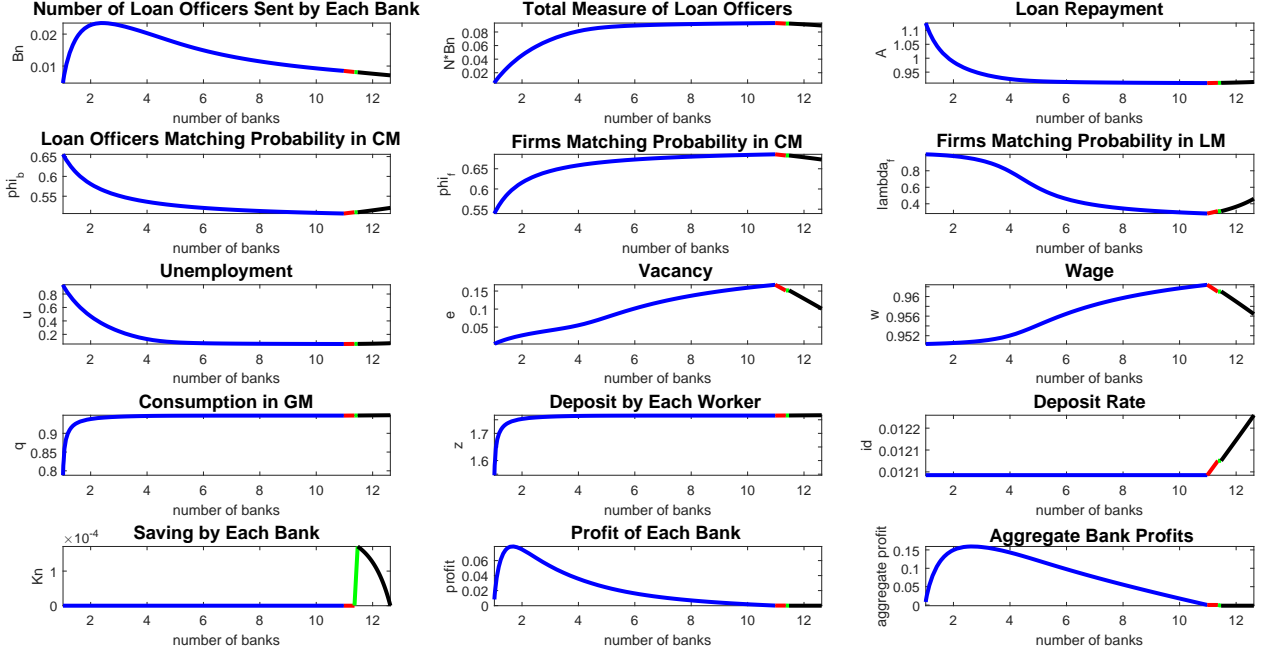


Figure 4: Equilibrium Variables

baseline and three alternative specifications. Welfare in our model is defined as

$$\mathcal{W} = \frac{1}{1 - \beta} \{ (1 - u)y + \alpha^h [u(q) - c(q)] + ul - v\gamma - Ek - \kappa B - N\kappa_f - \rho z\kappa_d \}, \quad (38)$$

which includes payoff for all workers and excludes all costs. Figure 5 displays a hump-shape relationship between the number of banks and welfare. Welfare first increases and then decreases as N rises. This result is robust across all specifications. In the baseline case, the highest welfare is achieved at $N = 4.82$.

The intuition for these results is the following: as in a standard Cournot model, oligopolistic competition stimulates quantities supplied. In our model, having more banks induces more intense competition and more loans being issued in aggregate. As a result, more firms receive funding to create vacancies and more workers find jobs. In addition, competition leads to higher deposit rates, which attract more deposits from workers and result in more output being traded in DM. All of these elements contribute to a positive effect on welfare.

However, as N increases the total banking costs (which includes the costs of sending

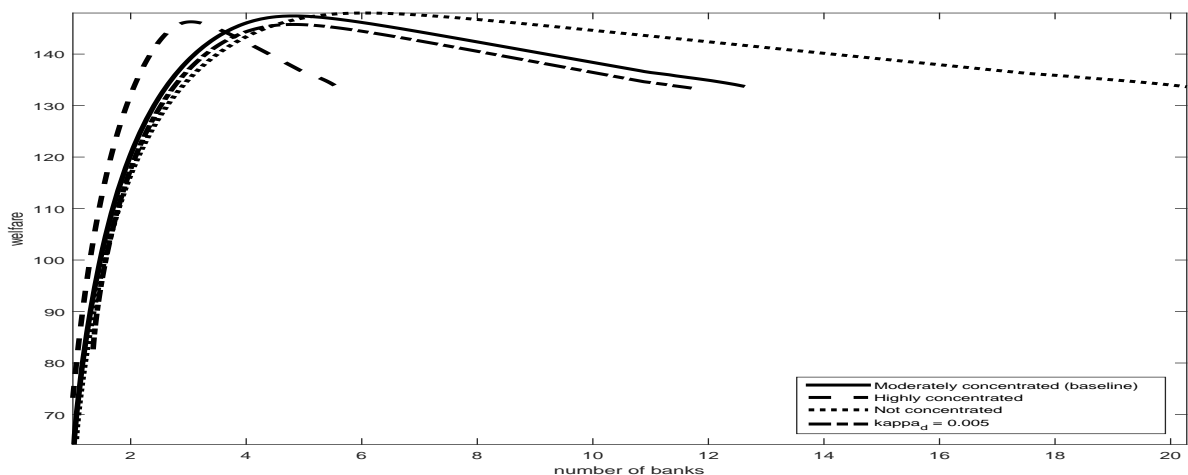


Figure 5: Welfare

bankers to the CM, κB , fixed costs, $N\kappa_f$, and variable cost of handling demand deposits, $\rho z\kappa_d$) also increase. More banks cause congestion in CM and reduce the probability of matching for each loan officer. Moreover, the increase in N implies that each bank receive a smaller share of deposits, banks must offer higher rates to attract deposit, which tightens banks' liquidity constraints. Lastly, as competition becomes intensified, banks have to start saving for themselves even though the saving is subject to inflation tax. All of these factors contribute to a negative effect on welfare as N rises.

Overall, when the number of banks is small, the positive welfare effects dominate because q increases and u decreases sharply with competition, reflecting the upward-sloping section in Figure 5. Once N goes beyond a certain threshold N^* (4.82 in the baseline case), the increase in q and decrease in u slow down and thus the negative effects of competition dominate. The relationship between welfare and N becomes negative.

Unlike a Cournot model in which competition always improves welfare, in our model, banking sector competition improves welfare only to a certain extent. This is a unique result from studying an oligopolistic banking sector. Oligopolistic banks must deal with the problem that how much liquidity it can “generate” for (the producer side of) the economy is limited by how much liquidity it can gather from (the consumer side of) the same economy. This is essentially a macroeconomic effect of Cournot competition. In our economy, it is socially optimal to have a small, yet greater than one, number of banks.

We also study how inflation affects the economy in the baseline model. Inflation is a

tax on cash-intensive activities. Thus, a higher inflation induces workers to economize on real balances, which adversely affects trading activities in DM. The lower trading surplus in DM discourages firms from entering the LM and thus reduces employment. Figures B2 and B3 in the Appendix suggest that inflation has the usual negative impact on welfare and economic activities in our economy.

It is a well-known fact that the banking industry is a heavily-regulated industry due to the important roles played by banks in the economy. In the U.S. banking system, banks must be chartered by either the state or federal government. Our results imply that bank charters can be used as a policy instrument to improve welfare by regulating the number of banks.⁹

5.2.2 Equilibrium with an Endogenous N

While the banking industry is certainly not characterized by free entry in practice, both banking profitability and regulations are important factors that influence creation of new bank charters. In this subsection, we consider how banking profitability affects bank entry. In this way, we endogenize the number of banks to study the effect of inflation on welfare. Note that the free-entry condition by banks indicates that the equilibrium number of banks N is determined by $\Pi_{bn}(N) = 0$. We use the calibrated parameter values for the baseline model to study how welfare vary with inflation, where annual inflation ranges from -2.773% to 20% .¹⁰

Figure 6 plots how $\Pi_{bn}(N)$ varies with N for different inflation rates.¹¹ The intersection of the profit function and the zero-profit condition determines the equilibrium N . Interestingly, a finite number of banks exists in equilibrium and inflation can change the nature of banking equilibrium. Intuitively, inflation makes workers deposit and consume less in DM. The lower trading surplus in DM reduces firms' incentive to enter the CM, which in turn drives down bank's profits. Meanwhile, inflation also makes the bank's

⁹The model in Huang (2019) has a similar policy implication, however, the mechanisms behind our results are different.

¹⁰As mentioned earlier, we focus on equilibrium with $i_d \geq 0$. According to condition (19), at the Friedman rule (i.e., $1/\rho = 1 + \pi = \beta = 0.992$, which yields a zero nominal interest rate), no banking equilibrium with $i_d \geq 0$ exists. Thus, the lower bound on $1 + \pi$ that allows a banking equilibrium with $i_d \geq 0$ to exist is given by $1 + \pi = \beta/(1 - \kappa_d) = 0.992/0.999$. In other words, annual inflation must be greater than or equal to $(1 - 7.007 \times 10^{-3})^4 - 1 = -2.773\%$.

¹¹We also report the relationship between inflation and equilibrium number of banks in Figure B4 in the Appendix. In addition, Figure B5 plots how a set of endogenous variables respond to inflation.

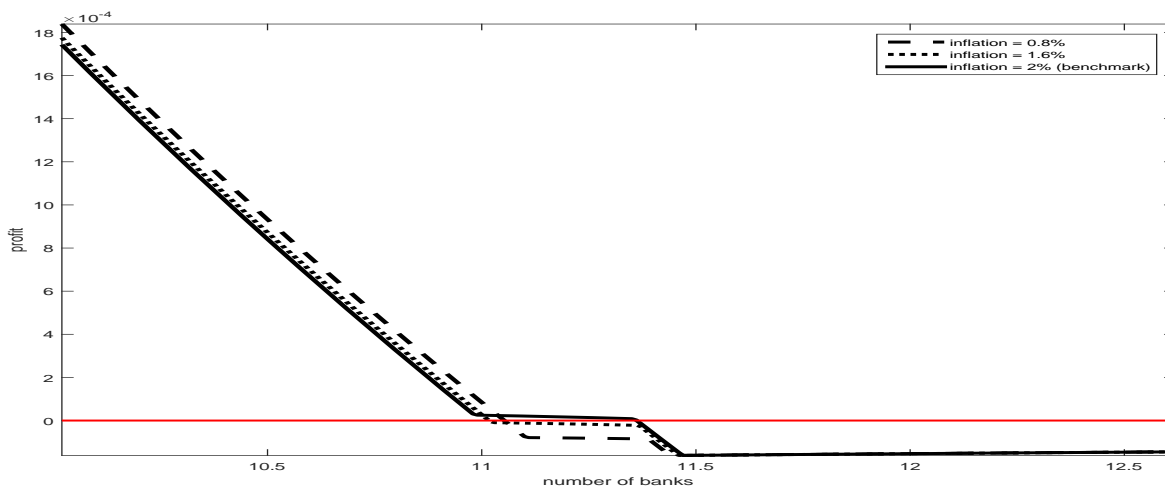


Figure 6: Inflation and Bank Profits

lending constraint more likely to bind and thus the bank's own savings become more necessary as N increases. Thus, as inflation increases, the profit function shifts down and equilibrium switches types. The economy is in a type I equilibrium when inflation is low but switches to a type II and then further to a type III equilibrium as inflation increases.

Figure 7 plots the relationship between inflation and welfare – a hump-shape relationship. In addition to the standard negative effect of inflation on welfare that exists in most monetary models, in our model, inflation can actually play a welfare-improving role at low inflation rates. This novel effect comes from the fact that equilibrium number of banks is inefficiently high.¹² Inflation reduces banks expected profit, which leads to less entries of banks at low inflation rates. This reduction in the number of banks is welfare-improving as it eases congestion in CM. However, as inflation goes beyond the optimal level (1.53% in our model), the negative effects of inflation start dominating. Any further increase in inflation reduces welfare. Our results shed new light on the conventional wisdom that moderate inflation may serve to "grease the wheels" of the economy through modeling an oligopolistic banking sector.

¹²With baseline parameters, at $\pi = -2.773\%$, (lower bound on π that ensures existence of a banking equilibrium with $i_d \geq 0$), Figure B4 in the Appendix indicates that in equilibrium $N = 11.216$. However, Figure B6 shows that welfare is maximized at $N = 4.42$.

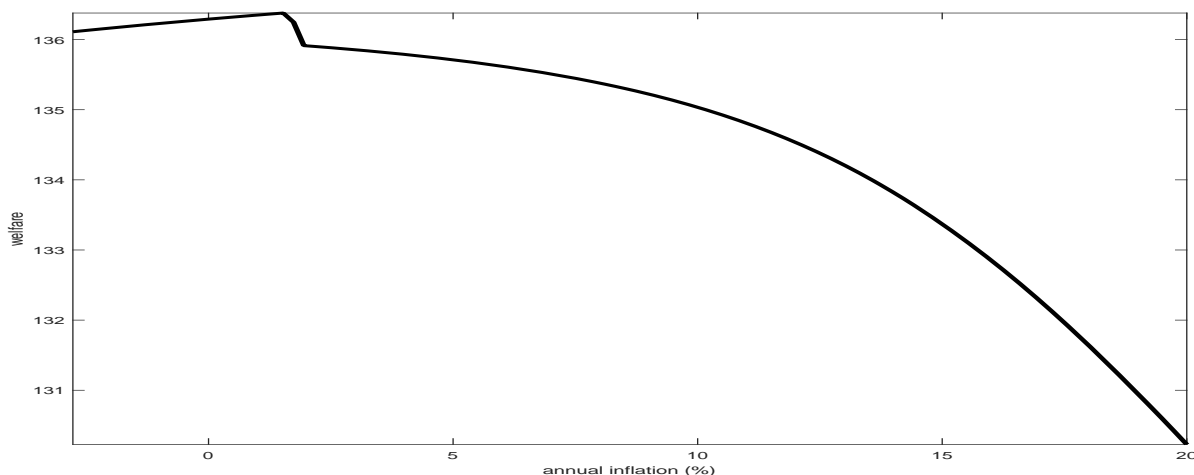


Figure 7: Inflation and Welfare

6 Conclusion

We have constructed a tractable macroeconomic model of money and banking to study the behavior and economic impact of oligopolistic banks. Our model has three key features: (i) banks as oligopolists, (ii) liquidity constraints for banks that arise from timing mismatch of cashflows, and (iii) frictions in credit, labor and goods markets.

We find that it is optimal to have a small, yet greater than one, number of banks. That is, an oligopolistic banking sector can be welfare-maximizing. When the number of banks is low and banks are not liquidity constrained, bank competition improves welfare. However, as more banks crowd the credit market, it becomes more difficult to make loans in the sense that the rate at which loans are issued declines. This decrease in lending opportunities and the rise in lending costs eventually dominate and make welfare decrease with the number of banks. Moreover, each bank will become liquidity constrained with more banks competing for funds in the deposit market, which will further strengthen the negative impact on welfare. With the number of banks being endogenous, we find that moderate inflation can improve welfare as it eases congestion in the credit market and helps to improve the efficiency of banking sector.

References

- [1] Andolfatto, D. (1996). “Business Cycles and Labor-Market Search.” *American Economic Review* 86, 112-132.
- [2] Berentsen A., G. Camera and C. Waller (2007). “Money, Credit and Banking.” *Journal of Economic Theory* 135, 171-195.
- [3] Berentsen, A., Menzio, G. and Wright, R. (2011). “Inflation and Unemployment in the Long Run.” *American Economic Review* 101, 371-98.
- [4] Bethune Z., G. Rocheteau G. and P. Rupert (2015). “Aggregate Unemployment and Household Unsecured Debt.” *Review of Economic Dynamics* 18, 77-100.
- [5] Black, L. and M. Kowalik (2017). “The Changing Role of Small Banks in Small Business Lending.” Manuscript.
- [6] Burdett, K. and K. Judd (1983). “Equilibrium Price Dispersion.” *Econometrica* 51(4), 955-969.
- [7] Cavalcanti, R. and N. Wallace (1999). “Inside and Outside Money as Alternative Media of Exchange.” *Journal of Money, Credit, and Banking* 31(3), 443-457.
- [8] Corbae, D. and P. D’Erasmus (2013). “A Quantitative Model of Banking Industry Dynamics.” Manuscript.
- [9] Davis, S. J., R. J. Faberman and J. Haltiwanger (2006). “The Flow Approach to Labor Markets: New Data Sources and Micro-macro Links.” *Journal of Economic Perspectives* 20, 3–26.
- [10] Diamond D. and P. Dybvig (1983). “Bank runs, Deposit Insurance, and Liquidity.” *Journal of Political Economy* 91, 401–419.
- [11] Faig, M. and B. Jerez (2005). “A Theory of Commerce.” *Journal of Economic Theory* 122, 60-99.
- [12] Freixas, X. and J.C. Rochet (2008). “Microeconomics of Banking.” Cambridge: MIT Press.

- [13] Gertler, M. and N. Kiyotaki (2010). “Financial Intermediation and Credit Policy in Business Cycle Analysis.” *Handbook of Monetary Economics*, Volume 3A, Elsevier, 547-599.
- [14] Gertler, M. and N. Kiyotaki (2015). “Banking, Liquidity and Bank Runs in an Infinite Horizon Economy.” *American Economic Review* 105, 2011-2043.
- [15] Gertler, M., N. Kiyotaki and A. Prestipino (2016). “Wholesale Banking and Bank Runs in Macroeconomic Modelling of Financial Crises.” *Handbook of Macroeconomics* Volume 2, Elsevier, 1345-1425.
- [16] Gorton G. and A. Winton (2003). “Financial Intermediation.” *Handbook of the Economics of Finance*, Volume 1, Elsevier, 431-552.
- [17] Hagedorn, M. and I. Manovskii (2008). “The Cyclical Behavior of Equilibrium Unemployment and Vacancies Revisited.” *American Economic Review* 98, 1692-1706.
- [18] He, P., L. Huang and R. Wright (2005). “Money and Banking in Search Equilibrium.” *International Economic Review* 46, 637-70.
- [19] He, P., L. Huang and R. Wright (2008). “Money, Banking and Monetary Policy.” *Journal of Monetary Economics* 55, 1013-1024.
- [20] Huang K. (2019). “On the Number and Size of Banks: Efficiency and Equilibrium.” Manuscript.
- [21] Kalai, E. (1977). “Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons.” *Econometrica* 45, 1623–1630.
- [22] Head, A., T. Kam, T., and L. Man (2020). “Equilibrium Imperfect Competition and Banking: Interest Pass Through and Optimal Policy.” Manuscript.
- [23] Lagos, R. and R. Wright (2005). “A Unified Framework for Monetary Theory and Policy Analysis.” *Journal of Political Economy* 113, 463-484.
- [24] Mortensen, D. and C. Pissarides (1994). “Job Creation and Job Destruction in the Theory of Unemployment.” *Review of Economic Studies* 61, 397-415.

- [25] Petrosky-Nadeau, N. and E. Wasmer (2013). “The Cyclical Volatility of Labor Markets under Frictional Financial Markets.” *American Economic Journal: Macroeconomics* 5, 193-221.
- [26] Petrosky-Nadeau, N. and E. Wasmer (2015). “Macroeconomic Dynamics in a Model of Goods, Labor, and Credit Market Frictions.” *Journal of Monetary Economics* 72, 97-113.
- [27] Petrosky-Nadeau, N., L. Zhang, and L. Kuehn (2018). “Endogenous Disasters.” *American Economic Review* 108, 2212-45.
- [28] Shimer, R. (2005). “The Cyclical Behavior of Equilibrium Unemployment and Vacancies.” *American Economic Review* 95, 25-49.
- [29] Silva, J. and M. Toledo (2009). “Labor Turnover Costs and the Cyclical Behavior of Vacancies and Unemployment.” *Macroeconomic Dynamics* 13, 76-96.
- [30] Sun, H. (2007). “Aggregate Uncertainty, Money and Banking.” *Journal of Monetary Economics* 54, 1929-1948.
- [31] Sun, H. (2011). “Money, Markets and Dynamic Credit.” *Macroeconomic Dynamics* 15, 42-61.
- [32] Wasmer, E. and P. Weil (2004). “The Macroeconomics of Credit and Labor Market Imperfections.” *American Economic Review* 94, 944-963.
- [33] Williamson, S. (2004). “Limited participation, Private money, and Credit in a Spatial Model of Money.” *Economic Theory* 24, 857-875.
- [34] Williamson, S. and R. Wright (2010a). “New Monetarist Economics: Models.” *Handbook of Macroeconomics*, Volume 3, Amsterdam: Elsevier, 25-96.
- [35] Williamson, S. and R. Wright (2010b). “New Monetarist Economics: Methods.” *Federal Reserve Bank of St. Louis Review* 92, 265-302.

A Proofs

A.1 Proof of Lemma 1

The Kalai bargaining problem in the DM is outlined in (24). If $d < z$, we can rearrange the surplus sharing rule to get (25). Then we can eliminate ρd in the objective function. The bargaining problem is reduced to $\max_q u(q) - \beta(1 + i_d)c(q)$. Therefore, the optimal choice of q solves $u'(q^*)/c'(q^*) = \beta(1 + i_d)$. Accordingly, the optimal choice of payment $d = z^*$, where z^* is given by $\beta\rho z^* = g(q^*)$. If $d = z$, q can be directly solved from (25). That is, $q(z) = g^{-1}(\beta\rho z)$.

A.2 Proof of Proposition 1

Consider condition (6), which defines optimal choice of z . Notice that $u'(q)\partial q/\partial z - \beta\rho(1 + i_d)\partial d/\partial z$ is zero for all $z \geq z^*$ by Lemma 1. Hence, $\beta\rho(1 + i_d) > 1$ implies that the problem of choosing z has no solution, because the objective function $\max_z \{\alpha^h[u(q) - \beta(1 + i_d)\rho d] - z[1 - \beta\rho(1 + i_d)]\}$ is strictly increasing for all $z \geq z^*$. This result means that any equilibrium must satisfy $\beta\rho(1 + i_d) \leq 1$. Given $\beta\rho(1 + i_d) \leq 1$, the objective function is non-increasing in z for $z \geq z^*$.

The slope of the objective function as $z \rightarrow z^*$ from below is proportional to $-[1 - \beta\rho(1 + i_d)] + \alpha^h[u'(q)\beta\rho/g'(q) - \beta\rho(1 + i_d)]$, where $u'(q)\beta\rho/g'(q) - \beta\rho(1 + i_d)$ is the worker's marginal gain from bringing an additional unit of deposit into the DM. For any $\mu \in (0, 1)$, unless $\beta\rho(1 + i_d) = 1$, the slope of the objective function as $z \rightarrow z^*$ is strictly negative, and therefore any solution must satisfy $z < z^*$. This completes the proof of the claim that $d = z$ as long as we are not in the extreme case of $\beta\rho(1 + i_d) = 1$.

A.3 Proof of Proposition 2

Consider the decisions of individual bank n . The first-order condition for $\hat{Z}_n > 0$ is given by

$$1 - \kappa_d + \varepsilon + \xi + \beta\hat{\mathcal{V}}_3 = 0. \quad (39)$$

The first-order condition for $B_n > 0$ is

$$\left[\phi^b(B_n, \bar{B}_{-n}, E) + B_n \frac{\partial \phi^b}{\partial B_n} \right] \left[(1 - \lambda^f) \beta \hat{\mathcal{V}}_1 + \lambda^f \beta \hat{\mathcal{V}}_2 - \gamma(1 + \varepsilon + \xi) \right] = \kappa(1 + \varepsilon + \xi). \quad (40)$$

The first-order condition for \hat{K}_n is

$$\beta \hat{\mathcal{V}}_4 - \xi - 1 \leq 0, \quad \hat{K}_n \geq 0. \quad (41)$$

Moreover, the Envelope Theorem yields

$$\mathcal{V}_1 = -\gamma(1 + \varepsilon + \xi) + \beta \left[(1 - \lambda^f) \hat{\mathcal{V}}_1 + \lambda^f \hat{\mathcal{V}}_2 \right] \quad (42)$$

$$\mathcal{V}_2 = A(1 + \xi) + \beta(1 - s) \hat{\mathcal{V}}_2 \quad (43)$$

$$\mathcal{V}_3 = -\rho(1 + i_d)(1 + \varepsilon + \xi) \quad (44)$$

$$\mathcal{V}_4 = \rho(1 + \varepsilon + \xi) \quad (45)$$

Equations (39) and (44) imply optimal condition for \hat{Z}_n is $1 + \varepsilon + \xi - \kappa_d - \beta\rho(1 + i_d)(1 + \varepsilon + \xi) = 0$. Thus,

$$\beta\rho(1 + i_d) = 1 - \frac{\kappa_d}{1 + \varepsilon + \xi} < 1, \quad (46)$$

given $\kappa_d > 0$. Therefore, the money constraint must be binding for all workers, i.e., $d = z$.

Equations (40), (42) and (43) imply the optimal condition for B_n is (18). Furthermore, (41) and (45) imply the optimal condition for \hat{K}_n is

$$\beta\rho(1 + \varepsilon + \xi) - (1 + \xi) \leq 0, \quad \hat{K}_n \geq 0. \quad (47)$$

We now describe possible equilibria. First, consider $\varepsilon = 0$. Then condition (47) yields $\beta\rho(1 + \xi) - (1 + \xi) < 0$, given that $\beta\rho < 1$. Thus, $\varepsilon = 0$ implies $\hat{K}_n = 0$. That is, a non-binding lending constraint (12) must mean that the bank does not carry a positive balance of \hat{K}_n . There are two possible cases. If the withdrawal constraint does not bind, neither constraint binds and $\hat{K}_n = 0$. The equilibrium is type I in which (B, E, u, v, A) are solved from (18), (22), (28), (29), and (36). If the withdrawal constraint binds, $\xi > 0$ and the equilibrium is type V. The values of (B, E, u, v, A, ξ) are determined by (13) with equality, (18), (22), (28), (29), and (36).

Second, consider $\varepsilon > 0$ and $\hat{K}_n = 0$. The lending constraint binds and banks do not save yet. The binding lending constraint (12) implies,

$$[1 - \rho(1 + i_d)(1 - \alpha^h)]z + \rho NK_n - \gamma[g_1(A) + B\phi^b(B, E)] - \kappa B - N\kappa_f = 0. \quad (48)$$

There are two cases to consider depending on whether the withdrawal constraint binds. If the withdrawal constraint does not bind, the equilibrium is type II. The values of (B, E, u, v, A, ξ) are solved from (18), (22), (28), (29), (36) and (48). If the withdrawal constraint binds, the binding lending constraint (12) implies that (13) can be simplified into

$$ag_{n2}(a) - D_n = \hat{K}_n. \quad (49)$$

The equilibrium is type VI. The values of $(B, E, u, v, A, \varepsilon, \xi)$ are solved from (18), (22), (28), (29), (36), (48) and (49).

Lastly, consider $\varepsilon > 0$ and $\hat{K}_n > 0$. The lending constraint binds and banks save. Then (47) yields (20). Combined with (46), we have $(1 + \varepsilon + \xi - \kappa_d)/(1 + i_d) = 1 + \xi$. Thus, $\varepsilon = (1 + \xi)i_d + \kappa_d > 0$. That is, lending constraint (12) must be binding as long as the bank holds a positive balance of \hat{K}_n . There are two possible cases here, depending on whether constraint (13) binds or not. When the constraint (13) does not bind, $\xi = 0$. The equilibrium is type III. The values of $(B, E, u, v, A, \varepsilon, \hat{K}_n)$ are determined by (18), (22), (28), (29), (36), (48) and (20). When the constraint (13) binds, both constraints bind and $\hat{K}_n > 0$. The equilibrium is type IV. The values of $(B, E, u, v, A, \varepsilon, \xi, \hat{K}_n)$ are determined by (18), (20), (22), (28), (29), (36) (48) and (49).

A.4 Welfare

Welfare in our model is defined as

$$\mathcal{W} = (1 - u)[\alpha^h W_e^h(0) + (1 - \alpha^h)W_e^h(\rho z)] + u[\alpha^h W_u^h(0) + (1 - \alpha^h)W_u^h(\rho z)] - \frac{Ek}{1 - \beta}.$$

Note that $\partial W_e^h(\rho z)/\partial(\rho z) = 1 + i_d$. Hence, $W_e^h(\rho z) = (1 + i_d)\rho z + W_e^h(0)$ and $W_u^h(\rho z) = (1 + i_d)\rho z + W_u^h(0)$. From (5) we know that $W_e^h(0) - W_u^h(0) = [w - (b + \ell)]/[1 - \beta(1 - s - \lambda^h)]$.

Therefore, $\mathcal{W} = W_u^h(0) + (1-u)[w - (b+\ell)]/[1 - \beta(1-s-\lambda^h)] + (1-\alpha^h)(1+i_d)\rho z - (Ek)/(1-\beta)$.
We know that

$$W_u^h(0) = \frac{b + \ell + \frac{\beta\lambda^h w}{1-\beta(1-s)} + [1 + \frac{\beta\lambda^h}{1-\beta(1-s)}]\{\Pi + \rho\mathbf{f}_{b,-1} - T + \alpha^h u(q) - z[1 - \beta\rho(1+i_d)(1-\alpha^h)]\}}{1 - \frac{\beta^2\lambda^h s}{1-\beta(1-s)} - \beta(1-\lambda^h)}.$$

Together with the facts that $\Pi = \alpha^f[\rho d - c(q)] - w - A$, $\mathbf{f}_{b,-1} = n\Pi_{bn} = n\rho K_n + (1-\alpha^h)z[1 - (1+i_d)\rho] - z\kappa_d - \gamma[g_1(A) + B\phi^b(B, E)] - \kappa B - n\kappa_f + Ag_2(A) - nK_n$, $T = bu - (1-\rho)M/p$, and the money-market clearing condition $M/p = Ag_2(A) + n\rho K_n + z[1 - (1-\alpha^h)(1+i_d)\rho] - \gamma[g_1(A) + B\phi^b(B, E)] - \kappa B - n\kappa_f$, the welfare function can be simplified into $(1-\beta)\mathcal{W} = (1-u)y + \alpha^h[u(q) - c(q)] + u\ell - v\gamma - Ek - \kappa B - N\kappa_f - \rho z\kappa_d$.

B Additional Figures

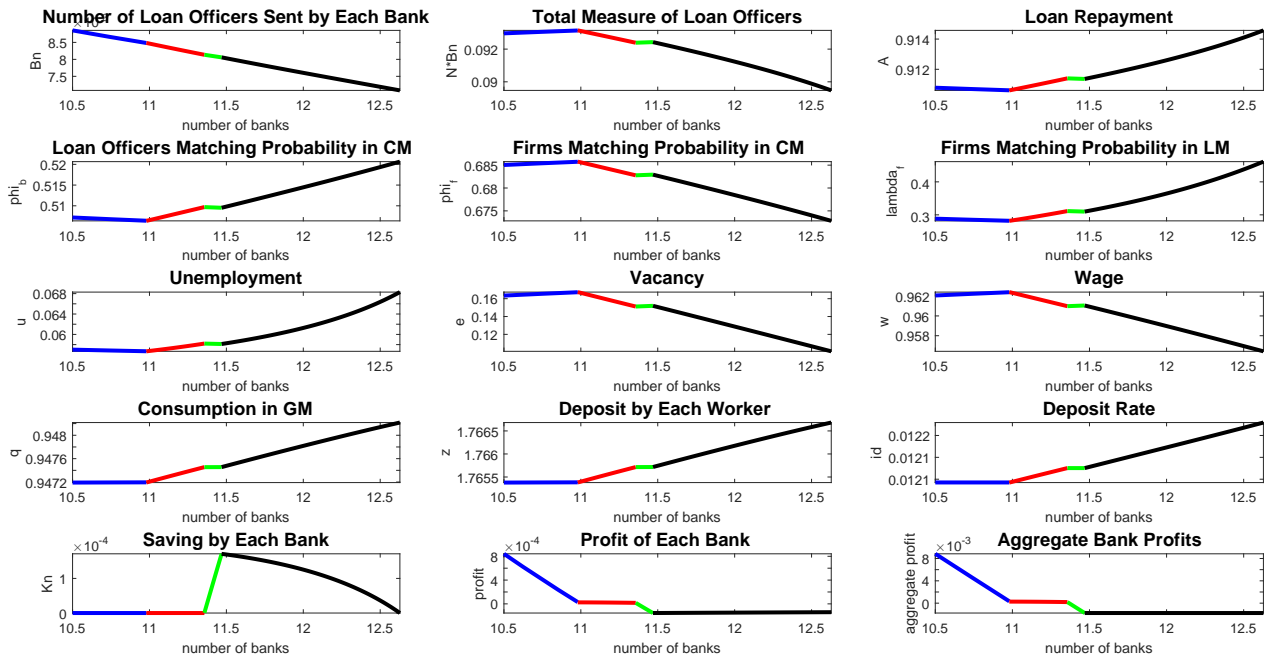


Figure B1: Equilibrium Variables (Partial View)

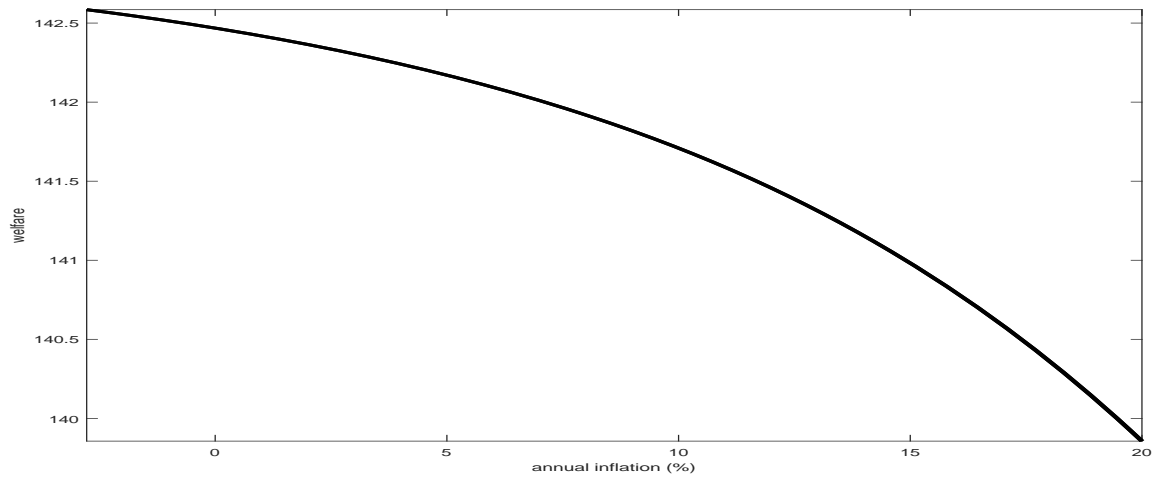


Figure B2: Inflation and Welfare ($N = 8$)

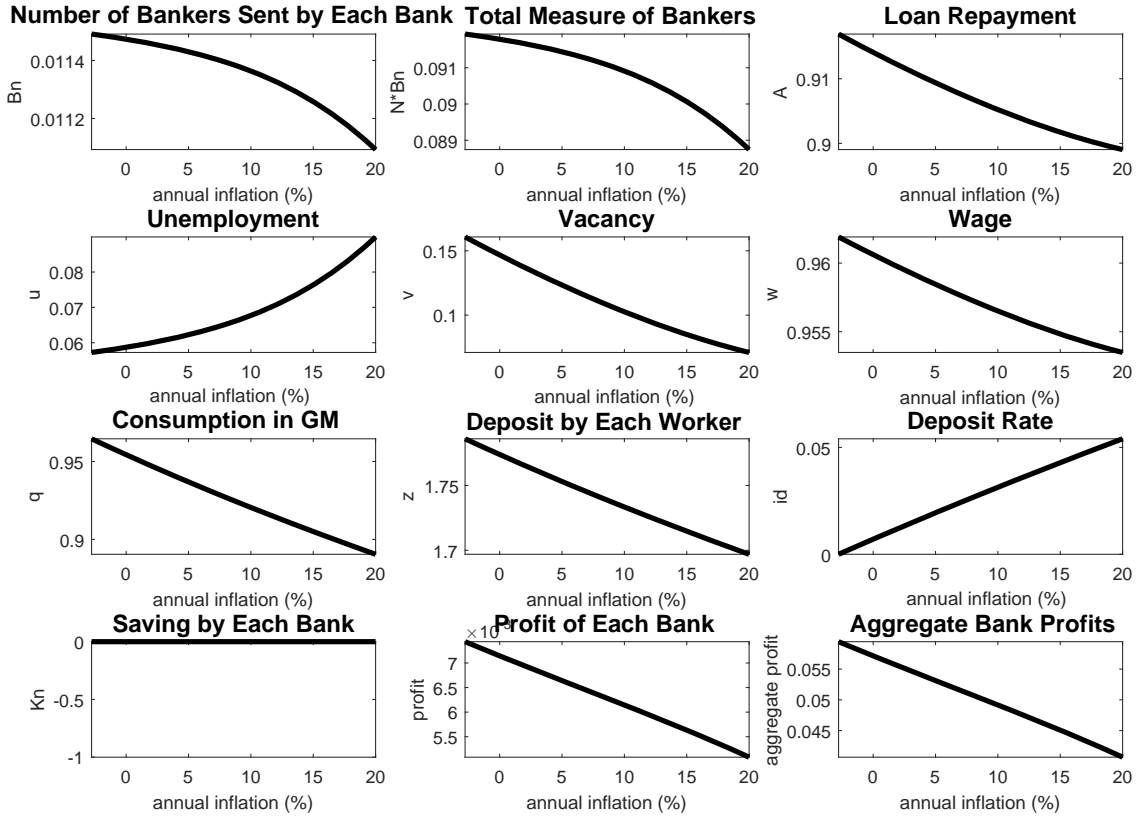


Figure B3: Inflation and Equilibrium Variables ($N = 8$)

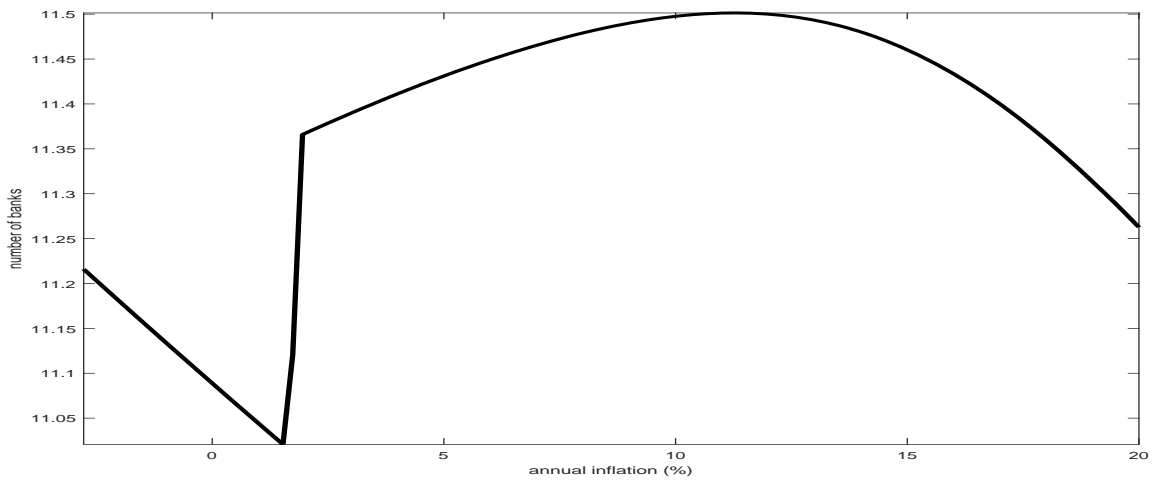


Figure B4: Inflation and Equilibrium Number of Banks (N)

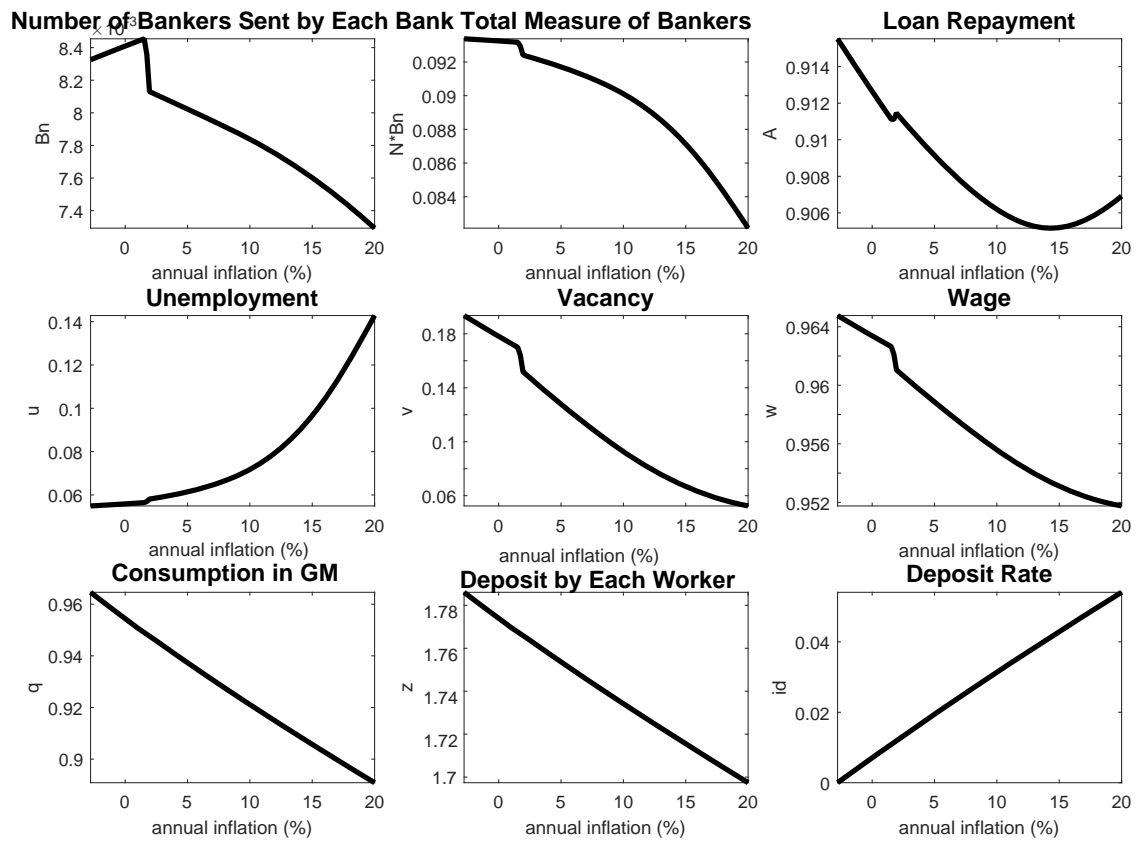


Figure B5: Inflation and Equilibrium Variables (Endogenous N)

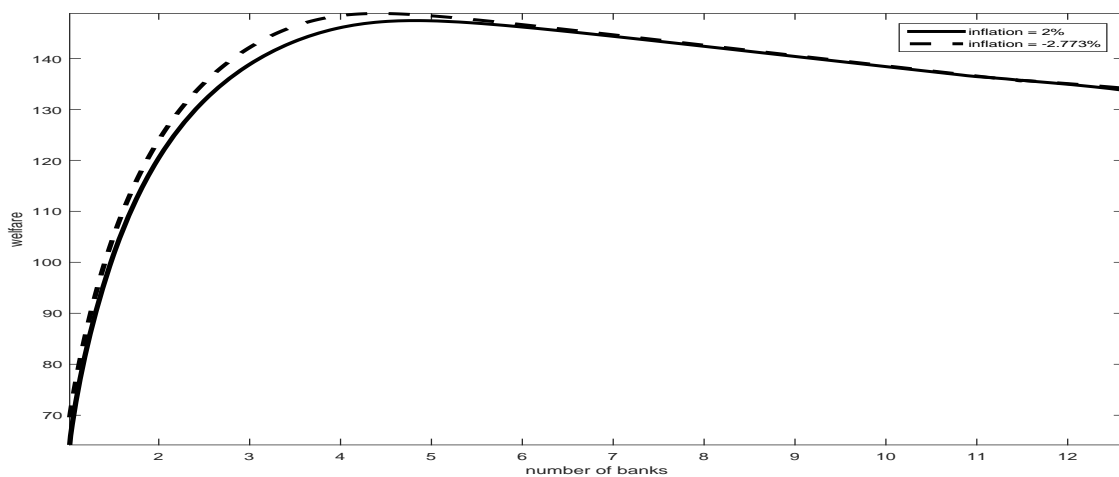


Figure B6: Welfare at Two Different Inflation Levels