

Banks' Wealth, Banks' Creation of Money, and Central Banking*

Tianxi Wang
University of Essex

Banks are special in that their liabilities are widely accepted as a means of payment, thereby often needed by real sectors to obtain resources. This paper studies this interaction between the banking sector and real sectors on competitive markets and the policy response of the central bank to market inefficiency, which is determined by the aggregate wealth of banks. In the circumstance of a credit crunch, the central bank improves efficiency by allowing banks to borrow its fiat money at zero interest up to a limit. This policy bears the flavor of quantitative easing policies (QE). It produces real effects in the absence of surprises and nominal rigidity. The mechanism in which it works depends on a difference in nature between bank-created money and fiat money. Furthermore, this policy, while expanding the money supply, induces deflation under the positive productivity shock. Lastly, this paper explains when interest rate policy and capital adequacy regulation are among the optimal policies within a unified model.

JEL Codes: E5, E65, G21.

“The community cannot get rid of its currency supply... The ‘hot potato’ analogy truly applies. For bank-created money, however, there is an economic mechanism of extinction as well as creation, contraction as well as expansion.”

James Tobin (1963)

*I am grateful to John Moore and Santiago Oliveros for the generously long discussion they gave of the paper. I thank Amil Dasgupta, Sudipto Dasgupta, Tore Ellingsen, Xuewen Liu, Stephan Niemann, and seminar participants at the University of Essex, University of Edinburgh, Zhejiang University, SHUFE, the Growth and Policy conference at Durham University, HKUST, the “Systemic Risk, Financial Markets and the Post-Crisis Economy” conference at the University of Nottingham, CICF 2013, the Econometrica Society World Congress 2015, and ETH Zurich for their helpful comments.

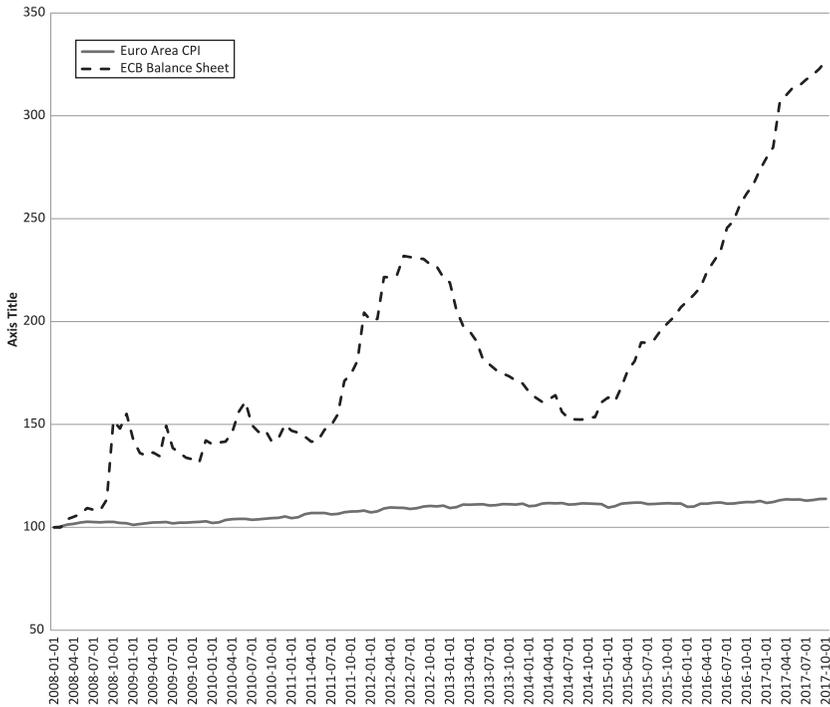
1. Introduction

Commercial banks are special in that their liability, especially that in the form of demand deposit, is *widely* accepted as a means of payment,¹ whereas rarely so are the liabilities of non-bank firms or households. Due to this difference, real-sector firms often need to borrow a bank's liability (usually called *money* in everyday language) as a means of paying for resources that they want. Banks' decisions on the price and quantity of money in lending and the competition between them, therefore, have a profound impact on the economic activity of real sectors, which, inversely, affects the decisions of and competition between banks. Furthermore, often, the central bank's policy produces effects by affecting banks' lending decisions and is based on its ability to create an alternative means of payment, namely fiat money. While money creation by banks is well known and widely introduced in macroeconomic textbooks, these interactions between real sectors, the banking sector, and the central bank in relation to means of payment have not been studied much yet,² a relationship which this paper makes an attempt to understand. Specifically, it presents a general equilibrium analysis of money creation by banks and of how certain policies of the central bank improve efficiency over the market allocations. This analysis

¹Often lay people perceive that depositing cash is to let the bank store the money deposited. However, this perception is wrong. Depositing is an exchange of cash to the bank's liability, which is what the depositor owns with the bank account and, when he makes a purchase with the account, is what he uses for the payment.

²The interaction between the banking sector and real sectors is studied in the literature on financial intermediation, where money creation by banks is not a concern; see Gorton and Winton (2003) for a survey. A strand of literature uses search-matching frameworks—see, e.g., Cavalcanti, Erosa, and Temzalides (1999) and Williamson (1999)—to examine how and when certain privately issued claims circulate as a means of payment, but has difficulty accommodating banks' decision on the price and scale of lending and competition between them. Lastly, New Keynesian literature and the literature that uses frameworks of cash in advance (CIA) both study monetary policy. However, the former abstracts banks completely, while the latter, when concerned with money creation by banks, lets the scale of banks' lending be pinned down either by a binding reserve constraint (see, e.g., Goodfriend and McCallum 2007) or by an exogenous rule of holding excessive reserves (see, e.g., Chen 2018 and Mishkin 2016), which leaves little role for banks' own decisions.

Figure 1. The Balance Sheet of the European Central Bank (ECB) and the Consumer Price Index in the Euro Area over Jan. 1, 2008–Oct. 10, 2017 (with Jan. 1, 2008 = 100)



Source of the Data: [www.https://fred.stlouisfed.org/](https://fred.stlouisfed.org/).

Note: The three hikes in the ECB's balance sheet during the period are associated with only slight price increases, if anything at all.

results in new insights on the mechanism in which the quantitative easing policy (QE) works, based on which this paper offers an explanation for the observation that QE, while causing enormous monetary expansion, is associated with low inflation or even deflation pressure in certain economies. One of them is the euro area, as is illustrated in figure 1.

The model economy of the paper is populated by workers, entrepreneurs, and banks. Workers can produce the consumption good, corn, in autarky, or work for entrepreneurs. The specialty of banks is

modeled with the assumption that workers accept banks' promises to pay, but not entrepreneurs', as a means of wage payment. At date 0, therefore, entrepreneurs first borrow banks' promises to pay—more specifically, notes that read like “X bank promises to pay the bearer 10 kilograms of corn tomorrow”—and then use these notes to hire workers. As a result, entrepreneurs owe a debt to the lender banks and banks owe a debt to the workers. At date 1, entrepreneurs produce corn and use it to settle their debts to the banks. Banks then use this repayment and their own corn, which is stored over time and represents their wealth, to redeem their notes from the workers by fulfilling the promises written on the notes.

In this economy, the real resources are workers' labor and entrepreneurs' capital, and efficiency is measured with the number of workers that entrepreneurs hire. Banks matter, however, because entrepreneurs need to borrow banks' liabilities to hire workers. How much banks lend in terms of real value determines how many workers entrepreneurs hire and hence economic efficiency. The aggregate lending of banks is in turn determined by competition between them. What banks supply is a means of payment, which is a homogeneous good. They thus engage in Bertrand competition. Moreover, what a bank lends is its liability. Hence the scale of its lending is anchored to its wealth by a borrowing constraint.³ Put together, banks are engaged in Bertrand competition with a limited capacity à la Kreps and Scheinkman (1983). If the borrowing constraint is binding, banks' aggregate wealth determines the quantity of money supplied to entrepreneurs, and hence efficiency. If this wealth is below a threshold, the money supply is inadequate and so is the number of workers that entrepreneurs hire.

This problem of meager bank wealth depressing economic activity has been diagnosed in many studies such as Gertler and Kiyotaki (2010) and Holmstrom and Tirole (1997). However, these studies are not concerned with means of payment. Hence they do not consider the possibility of a remedy with the central bank (CB) issuing fiat money. By contrast, in this paper the CB can offer a remedy by allowing banks to borrow its fiat money at zero interest up to a

³This borrowing constraint could be due to a moral hazard friction à la Gertler and Kiyotaki (2010) or Holmstrom and Tirole (1997), or due to the risk-shifting problem à la Jensen and Meckling (1976).

bound. We show that although the economy lasts for only two periods, in one equilibrium the fiat money circulates, flowing through banks to entrepreneurs, who thereby hire more workers. That is, the policy eases the constraints imposed upon their economic activity by an inadequate supply of bank-created money. Hence, the policy is called *quantitative easing policy*, or QE.

QE enables banks to raise lending capacity without violating the borrowing constraint, while backed by the same amount of wealth. It does so because fiat money is different in nature from bank-created money. A central bank never commits to redeem its fiat money at a specified value, whereas bank-created money is the bank's liability, which it commits to redeem at a pre-specified value and thus bears a real obligation of repayment.⁴ It is to this difference in nature that Tobin (1963) alludes above. As fiat money is not redeemable, its value freely adjusts with the state of the economy. In particular, its value falls in the event of the negative productivity shock, which means *inflation*. Inflation reduces the real value of banks' liability to the CB and thus reduces their borrowing constraint, giving them room to increase lending.⁵

While QE induces inflation in the event of the negative shock, it induces deflation in the event of the positive shock. This result might partly explain the aforementioned observation that in some economies QE causes a great monetary expansion on the one hand and is accompanied by or followed by lingering low inflation or even deflation pressure on the other hand, which is a puzzle if considered from the point of view of quantitative theory of money. A key difference in this paper is that the money created with QE is utilized to expand real economic activity, resulting in a rise in the output of goods, which keeps the price down. Indeed, all the QE-created money

⁴In the model economy, this fact is straightforward because banks use corn, the real good, to redeem their liability. Even if they redeem it with fiat money, which is typically the case in modern times, they still need to spend real resources to obtain it. Hence bank-created money still bears real obligations of repayment. For more discussion, see the remark in section 5.

⁵The point that monetary policy can help banks by affecting the real value of their liabilities is also considered by Diamond and Rajan (2006). The way in which fiat money circulates in the model economy is also in Allen and Gale (1998). However, these papers are not concerned with banks lending their liabilities to real sectors to be used as means of payment, nor with the inefficiency related to banks' wealth, nor with the quantitative easing policy.

is utilized this way if the scale of QE is below a threshold. If it is beyond it, the excess is not put into circulation, we also find. This finding partly explains another related observation, that in some economies a substantial fraction of the monetary bases created with QE is not lent out but stays in banks' reserve accounts.

This paper is in line with nascent literature that examines banks' specialty of their liability being accepted as a means of payment. Most closely related is Donaldson, Piacentino, and Thakor (2018). Both their paper and this one consider how banks' issuance of means of payment affects the real economy as well as policy implications. The two papers, however, have different focuses. Their paper explains how this specialty of banks is derived from their superior technology of warehousing, while this paper emphasizes the importance of banks' wealth for economic efficiency. Also the policy implications are different. Their paper shows that, contrary to the received wisdom, a higher central bank rate could raise bank lending, while this paper considers QE. Jakab and Kumhof (2015) describe in detail how banks create money with double bookkeeping. They focus on the quantitative implications of this facet of banking in a full-fledged dynamic stochastic framework, but are not concerned with economic efficiency and policy responses, on which this paper focuses.

Since the recent crisis, many studies have examined QE; see Gertler and Karadi (2011) and Gertler and Kiyotaki (2010), among others. Their diagnosis on the source of the problem is shared by this paper: banks' wealth is too low, causing inadequate lending.⁶ However, those studies model banks not as issuers of means of payment but as intermediaries of trading real goods. As a result, in those studies what the government issues must be backed by tax incomes and is essentially the sovereign debt, whereas what the central bank issues in this paper is purely nominal. Furthermore, the mechanism by which QE works is different. In those studies, it works by transferring wealth to banks, whereas in this paper it works by reducing the real value of banks' liability via inflation.

⁶This is also true in a study by Brunnermeier and Sannikov (2013), where money plays the role of a saving instrument, as in Samuelson (1958), rather than the role of means of payment.

Other cases of inefficiency in connection with private issuance of means of payment are considered by Hart and Zingales (2015), Monnet and Sanches (2015), and Stein (2012). In Hart and Zingales (2015) and Stein (2012) inefficiency is driven by fire-sale externalities. In Monnet and Sanches (2015) inefficiency arises because banks offer a return rate on the liability side that is lower than the inverse of time discount. Besides this difference in the source of inefficiency, none of those studies considers the circumstance of banks under-lending or the policy of QE, as this paper does.

The rest of the paper is organized as follows. Section 2 sets up the model. Section 3 analyzes a benchmark case where banks face no borrowing constraint and are thus subject to unfettered forces of competition. This analysis not only examines the bank competition in the purest form but also bears relevance to the historical banking. Section 4 introduces the borrowing constraint of banks. Section 5 studies QE, and section 6 studies interest rate policy and capital adequacy regulation. Section 7 concludes. All proofs are relegated to the appendix.

2. The Model

The economy has one storable good, corn, used as the numeraire, and lasts for two days. Contracting and production occur at $t = 0$, yielding and consumption at $t = 1$. There are N banks, N^2 entrepreneurs, and N^3 workers, where N is a large number (later, in section 5, the central bank will be introduced).⁷ Thus, banks are in perfect competition and each serves a large number of entrepreneurs; and there are more workers than entrepreneurs can hire. All agents are risk neutral and protected by limited liability.

Workers either produce w kilograms (kg) of corn in autarky or are hired by entrepreneurs, who each have h units of human or physical capital. If an entrepreneur hires L workers at $t = 0$, then his project yields at $t = 1$

$$y = \tilde{A}h^{1-\alpha}L^\alpha,$$

⁷A setting with the same feature is to be found in Wang (2015), where a continuum of $[0, 1]$ and $[0, 1] \times [0, 1]$ instead of N and N^2 is used. These two representations are equivalent.

where $0 < \alpha < 1$. Without losing any generality, normalize $h = 1$. Productivity, \tilde{A} , is subject to a common shock. At $t = 0$, it is common knowledge that the good state with $\tilde{A} = \bar{A}$ occurs with probability $q > 0$ and the bad state with $\tilde{A} = \underline{A}$ occurs with probability $1 - q > 0$. Let $A_e \equiv q\bar{A} + (1 - q)\underline{A}$ denote the mean. Assume the following:

$$0 < \underline{A} < A_e \alpha. \quad (1)$$

As there are more workers than can be hired by entrepreneurs, in equilibrium workers are indifferent in working for the latter or in autarky. Therefore they earn a real wage of w . This wage is independent of economic activity of the other sectors, which gives a convenience for exposition.

Banks each have G units of corn, where a unit is defined as N kg and used wherever banks are concerned. Banks supply no real resources for corn production. What makes banks relevant is due to the following assumption.

ASSUMPTION 1. Workers do not accept entrepreneurs' promise to pay but banks' as a means of wage payment.

This assumption captures the speciality of banks explicated in the Introduction. According to Kiyotaki and Moore (2001), this difference between banks and entrepreneurs arises because the former has stronger commitment power than the latter.

Due to this assumption, entrepreneurs cannot hire workers at $t = 0$ with a promise to pay them later at $t = 1$. To hire workers, entrepreneurs need to borrow banks' promise to pay them at $t = 1$. We assume that banks have no difficulty enforcing repayment from entrepreneurs and hence this borrowing is feasible. To fix the idea, suppose that banks' promises to pay are printed on notes. That is, a note issued by a bank reads that this bank promises to pay the bearer of this note with a certain quantity of corn at $t = 1$.⁸ This quantity is the note's *face value* or *par value*.

⁸In the model economy banks promise to pay the real good in redeeming their liability, whereas in modern times, they typically redeem it with fiat money, the central bank's issues. Other instances of claims on real goods used as a means of

ASSUMPTION 2. *The face values of banks' notes cannot be contingent on the realization of \tilde{A} .*

This assumption captures the observation that in real life a private security that serves as a means of payment—such as demand deposit, promissory notes, cheques, or trade credit—commonly bears fixed claims and is of debt, and is rarely a contingent claim.⁹

Besides the friction of payment, entrepreneurs lack commitment power in another dimension.

ASSUMPTION 3. *Entrepreneurs are unable to make commitment on the scale of their projects in terms of the number of workers they will hire.*

This friction is *real* in the sense that it is unrelated to means of payment. In the absence of the friction, a bank would be willing to lower the interest rate, denoted by r , to a borrower entrepreneur who commits to a smaller scale, because thereby his project delivers a higher average return rate. This, however, would give entrepreneurs an incentive to borrow from multiple banks, each in a small amount and thus at a favorable rate. The presence of the real friction, therefore, is justified if entrepreneurs cannot be prevented from doing so. The importance of its presence is that it engenders a wedge between the first- and second-best allocations, as will be shown.

Due to the friction, a bank posts a single interest rate r for loans of any size E rather than a menu of $r(E)$. A loan contract is represented by a profile of (E, r) : at $t = 0$ the entrepreneur borrows the bank's notes of overall face value E , and at $t = 1$ he is obliged to pay $E(1 + r)$ kg corn back to the bank.

The timing of events is as follows. At $t = 0$, each bank posts the aggregate face value of notes that it will issue, D , and the interest rate that it will charge, r . Observing all these offers, each entrepreneur then chooses one bank to go to and asks to borrow face value

payment include Hart and Zingales (2015) and Williamson (1999). This abstraction is harmless in the baseline model where inflation is not addressed. Where it is, as in section 5 below, the abstraction helps clarify the mechanism by which QE works.

⁹For the reason why that is the case, see Dang, Gorton, and Holmstrom (2012) and Gorton and Pennacchi (1990).

E of the bank's notes. If one bank's notes are over-demanded, only a fraction of the entrepreneurs have their demand met. Entrepreneurs then use borrowed notes to hire workers and start the production. Banks store their G units of corn.

At $t = 1$, entrepreneurs produce corn. They either repay $E(1+r)$ kg corn to the lender banks or default. In that case they give all their output of corn to the banks. After receiving repayments from the entrepreneurs, the sum total denoted by \tilde{Y} , banks redeem notes from workers. If $\tilde{Y} + G \geq D$ to a bank, its notes are redeemed at par. If $\tilde{Y} + G < D$, the bank defaults and the notes are redeemed at a fraction $(\tilde{Y} + G)/D$ of their par values. Finally, the agents consume the corn in their possession.

In anticipation of the possibility of default, at $t = 0$ a bank's notes are discounted with a factor of

$$\delta = E_{\tilde{A}} \left[\min \left(1, \frac{\tilde{Y} + G}{D} \right) \right]. \quad (2)$$

Moving on to the equilibrium analysis, we figure out two benchmark allocations.

2.1 *The First-Best and Second-Best Allocations*

Efficiency concerns the number of workers allocated to entrepreneurs. Define the first-best allocation as the number of workers that maximizes the social surplus of projects, which is $A_e L^\alpha - wL$ due to universal risk neutrality and the opportunity cost of labor being w . The first-best allocation is thus

$$L^{FB} = \left(\frac{A_e \alpha}{w} \right)^{\frac{1}{1-\alpha}}.$$

The second-best allocation is defined as the number of workers that entrepreneurs would hire in the competitive equilibrium if the friction of payment (in assumption 1) were absent but the real friction (in assumption 3) remained—that is, if entrepreneurs could hire workers with their own promise to pay, but their wage offer could not be conditional on the scale of their projects. The equilibrium allocation is as follows.

LEMMA 1. *The second-best number of workers that entrepreneurs hire is*

$$L^{SB} = \left(\frac{q\bar{A}\alpha + (1-q)\underline{A}}{w} \right)^{\frac{1}{1-\alpha}}.$$

Obviously, $L^{SB} > L^{FB}$. That is, the real friction induces the entrepreneurs to hire too many workers.¹⁰ This fact engenders a circumstance of banks over-lending, among remedies to which are interest rate policy and capital adequacy regulation, as will be shown in sections 4 and 6.

Below we first analyze the baseline model in which bank issuance is subject to no borrowing constraints nor any other restrictions. This analysis serves two purposes: (i) it studies the competitive equilibrium of money creation by banks in the purest form; (ii) it bears relevance to early periods of banking history.

3. The Least-Fettered Issuance

In the least-fettered issuance, banks *can* finance an unlimited quantity of assets by issuing promises to pay. The *equilibrium* quantity is of course limited, as examined below. We first consider the demand side of the market for banks' notes, then the supply side, and, finally, the meeting of the two.

3.1 The Demand Side of the Market for Notes

Consider a representative entrepreneur who borrows from a bank offering (D, r) . If he borrows notes of a face value E , then they are worth δE , where the discount factor δ , as will be shown, is a function of (D, r) . With these notes, he hires workers in a number of

$$L = \frac{\delta E}{w}, \tag{3}$$

because they earn real wage w . At $t = 1$, the entrepreneur either repays $E(1+r)$ of corn to the bank or defaults. Thus, his decision problem is

¹⁰For a more general analysis of this type of inefficiency, see Wang (2010).

$$\max_E E_{\tilde{A}}[\max(\tilde{A}L^\alpha - E(1+r), 0)], \text{ s.t. (3).}$$

LEMMA 2. For any (w, δ, r) , the solution to the above problem satisfies $\underline{A}L^\alpha < E(1+r)$. That is, entrepreneurs all default in the bad state.

This lemma is driven by the assumption that $\underline{A} < A_e\alpha$, which says that the negative shock is severe enough to knock entrepreneurs into default.

At the optimum, the demand of the representative entrepreneur for the notes is

$$E(\delta, r) = \left(\frac{\bar{A}\alpha}{1+r}\right)^{\frac{1}{1-\alpha}} \left(\frac{\delta}{w}\right)^{\frac{\alpha}{1-\alpha}} \quad (4)$$

and the number of workers that he hires and his profit are, respectively,

$$L(R) = \left(\frac{\bar{A}\alpha}{wR}\right)^{\frac{1}{1-\alpha}} \quad (5)$$

$$V(R) = q(1-\alpha) \left(\frac{\bar{A}^\frac{1}{\alpha}\alpha}{wR}\right)^{\frac{\alpha}{1-\alpha}}, \quad (6)$$

where

$$R \equiv \frac{1+r}{\delta}. \quad (7)$$

Note that, so defined, R is the *actual* interest rate and measures the cost of borrowing: to obtain a means of payment that is worth 1, the entrepreneur borrows notes of face value $1/\delta$, thus acquiring a debt of $(1+r)/\delta$. Naturally, his scale of hiring and his profit, both of which are real variables, are inversely related to the cost of borrowing, namely, R .

Recall that the efficiency concerns only the number of the workers hired by entrepreneurs, which depends only on the actual interest rate. The efficiency of market equilibrium is thus determined solely by the actual interest rate in equilibrium. Define R^{FB} (R^{SB}) as the

value of the actual interest rate at which entrepreneurs hire the first-best (second-best) number of workers, that is, $L(R) = L^{FB}$ (L^{SB}). With (5),

$$R^{FB} = \frac{\bar{A}}{A_e}$$

$$R^{SB} = \frac{\bar{A}\alpha}{q\bar{A}\alpha + (1 - q)\underline{A}}.$$

After banks all have posted (D, r) , each entrepreneur decides which bank to go to. In the equilibriums of this subgame, an entrepreneur gets the same expected profit, \widehat{V} , from any bank that attracts a number of entrepreneurs.¹¹ As entrepreneurs' profit depends only on the actual interest rate, define \widehat{R} by $V(\widehat{R}) = \widehat{V}$. Then, \widehat{R} is the actual interest rate that prevails on the notes market, conditional on banks' choices of (D, r) . Given that there is a large number of banks, any single bank is too small to affect \widehat{R} with its choice of (D, r) and takes it as given when making that choice.

3.2 The Supply Side of the Market for Notes

Consider a representative bank. To attract entrepreneurs, the bank's choice of (D, r) satisfies $V(R) \geq \widehat{V}$ or, equivalently, $(1 + r)/\delta \leq \widehat{R}$. While the interest rate r is directly chosen by the bank, the discount factor of its notes δ is determined by (D, r) . It depends on whether the bank ever defaults or not, that is, whether $D > \widetilde{Y} + G$ in some state at $t = 1$. In the good state, $\widetilde{Y} = \bar{Y} = D(1 + r)$ because the entrepreneurs do not default. As a result, the bank does not default. In the bad state, by lemma 2, all the entrepreneurs default and pass the output of their projects on to the bank, of which the aggregate value is given below.

¹¹By offering (D, r) , which determines δ , each bank chooses an actual interest rate $R = (1 + r)/\delta$. No entrepreneur goes to a bank offering $V(R) < \widehat{V}$ when he can get \widehat{V} elsewhere. On the other hand, if a bank offers $V(R) > \widehat{V}$, it induces over-demand for its notes (which is never optimal), so an entrepreneur coming to it is served with such a probability l that $l \cdot V(R) = \widehat{V}$.

LEMMA 3. *The aggregate value of the bank's loans in the bad state is*

$$\underline{Y} = \frac{A}{A\alpha} \times D(1+r).^{12} \quad (8)$$

In the bad state, the bank does not default if and only if $D \leq G + \underline{Y}$, which, with \underline{Y} given by (8), is equivalent to

$$D \cdot \left(1 - \frac{A(1+r)}{A\alpha}\right) \leq G. \quad (9)$$

On the left-hand side is the loss made by lending in the bad state, $D - \underline{Y}$. Thus, the inequality says that banks will stay solvent in the bad state if and only if their lending scale, D , is not too large relative to their wealth, G , so that the loss from loans can be absorbed by the wealth.

Substitute the value of \tilde{Y} given above into (2), and the discount factor of the bank's notes is determined by its choice of (D, r) via

$$\delta(D, r) = \left\{ \begin{array}{l} 1, \text{ if (9) is satisfied} \\ q \times 1 + (1-q) \times \left(\frac{G}{D} + \frac{A(1+r)}{A\alpha}\right), \text{ otherwise} \end{array} \right\}. \quad (10)$$

Now consider the representative bank's decision problem at $t = 0$. Taking into account the possibility of default, its economic profit with the choice of (D, r) , denoted by $\Pi(D, r)$, is $E_{\tilde{Y}} \max(G + \tilde{Y} - D, 0) - G$.

LEMMA 4.

$$\Pi(D, r) = D \left[\frac{1+r}{R^{SB}} - \delta \right]. \quad (11)$$

Therefore, the profit margin of lending—that is, the profit of lending out a note of face value 1—is $(1+r)/R^{SB} - \delta$. Intuitively, the present value of this note, which is part of the bank's liability, is δ , while the bank's revenue from such lending is $E_{\tilde{Y}} (\tilde{Y}/D) = (1+r)/R^{SB}$.

¹² $\underline{Y} < \bar{Y}$ because $A < A_e\alpha$ by assumption and $A_e\alpha < \bar{A}\alpha$.

The representative bank chooses (D, r) to maximize $\Pi(D, r)$ subject to the constraint that it can attract entrepreneurs, that is,

$$\frac{1+r}{\delta(D, r)} \leq \widehat{R}. \quad (12)$$

Given the scale of lending D , the bank wants to charge an interest rate as high as possible so long as it can attract entrepreneurs. Therefore, at the optimum, the constraint (12) is binding.¹³ Hence, the bank's optimal choice of r as a function of its choice of D , denoted by $r(D)$, is determined by

$$\frac{1+r}{\delta(D, r)} = \widehat{R},$$

and banks' problem becomes

$$\max_D D\delta(D, r(D)) \times \left[\frac{\widehat{R}}{R^{SB}} - 1 \right].$$

As $\delta > q$ always (because the bank always redeems its notes at par in the good state), the following proposition is self-evident.

PROPOSITION 1. *The solution to and the value of the representative bank's problem are*

- (i) if $\widehat{R} > R^{SB}$, then $D = \infty$ and $\Pi = \infty$;
- (ii) if $\widehat{R} = R^{SB}$, then $\Pi = 0$, and the bank is indifferent to any value of D , with $r = r(D)$;
- (iii) if $\widehat{R} < R^{SB}$, then lending makes a loss, and thus $D = 0$ and $\Pi = 0$.

As was shown, the profit margin of lending equals $\delta \left(\widehat{R}/R^{SB} - 1 \right)$ and is positive if and only if $\widehat{R} > R^{SB}$. Moreover, if the profit margin of lending is positive, banks obtain $\Pi = \infty$. That is because,

¹³Mathematically, given D , both $\frac{1+r}{R^{SB}} - \delta(D, r)$ and $\frac{1+r}{\delta(D, r)}$ increase with r .

despite their limited stocks of corn, they all have an unlimited lending capacity (i.e., $D = \infty$), due to two reasons: (i) what banks lend out is their promise to pay, essentially word of mouth, which they can infinitely supply; (ii) banks are subject to no constraints on the quantity of supply in the baseline model.

3.3 The Equilibrium: The Second-Best Allocation Attained

The prevailing actual interest rate, \widehat{R} , which plays the role of price, clears the market for notes in equilibrium. We define the symmetric equilibrium below and discuss other equilibriums later.

DEFINITION 1. A profile $(\{D, r, \delta, E\}; \widehat{R})$ forms an equilibrium if

- (i) given \widehat{R} , banks' choice of (D, r) is optimal and thus given in proposition 1. This choice determines $\delta = \delta(D, r)$ through (10);
- (ii) given that all banks offer the same (D, r, δ) , entrepreneurs go to each bank with the same probability and, by the law of large numbers, each bank receives N of them. Their demand for notes, E , is optimal, that is, $E = E(\delta, r)$ given by (4);
- (iii) the market clears: $D = E$.¹⁴

In any equilibrium, banks neither obtain an infinitely large profit nor abstain from lending, which, by proposition 1, is the case if and only if

$$\widehat{R} = R^{SB}.$$

Therefore, the real allocation in equilibrium is unique and conforms with the second-best allocation, the one that would arise if the friction of payment were absent and entrepreneurs could hire workers with their own promises to pay. Intuitively, what banks supply is a means of payment, a homogeneous good. Furthermore, in this case of unfettered issuance, they all have unlimited capacity. Therefore,

¹⁴Note that D is in the unit of N kg, while entrepreneurs' demand is denoted with the unit of kg.

they are in Bertrand competition, which annihilates their profit margin. As a result, entrepreneurs overcome the friction of payment at no costs and the real allocation is the one that would arise if the friction were absent, that is, the second-best allocation.

On the nominal side, however, there is indeterminacy. At $\widehat{R} = R^{SB}$, by proposition 1, the profit margin of lending is 0, and individual banks are indifferent to any quantity of issues, although in aggregation their issues exactly suffice for entrepreneurs to hire L^{SB} workers. This indeterminacy leads to a continuum of equilibria besides the symmetric one defined above. In the symmetric equilibrium, all banks issue the same quantity of notes and their notes are discounted at the same factor; therefore, one bank's notes are perfect substitutes for another's. In asymmetric equilibria, however, some banks issue more than others despite ex ante being identical, and see their notes more heavily discounted.

PROPOSITION 2.

- (i) *In any equilibrium, independent of banks' wealth G , $\widehat{R} = R^{SB}$, the profit margin of bank lending is 0, and the second-best allocation is attained.*
- (ii) *In any equilibrium, a fraction of banks default at $t = 1$ upon the realization of $\tilde{A} = \underline{A}$ if and only if banks' wealth is below a threshold G^* , where*

$$G^* = [(q\bar{A}\alpha + (1 - q)\underline{A}) - \underline{A}] \left(\frac{q\bar{A}\alpha + (1 - q)\underline{A}}{w} \right)^{\frac{\alpha}{1-\alpha}}.$$

An intuition for result (i) is given above. As for (ii), note that, given that banks are indifferent to any quantity of issues, the quantity of money circulated—that is, the aggregate bank lending—is determined by the demand side, namely entrepreneurs. Thus, the aggregate bank issues are fixed at a quantity exactly sufficient for entrepreneurs to hire the second-best number of workers. With this scale of lending, if banks' wealth, G , is sufficiently low—namely $G < G^*$ —then it is insufficient to absorb the loss incurred in the bad state and bank default occurs accordingly.

The baseline mode examined above, besides providing a simple framework to consider money creation by banks on competitive

markets, also bears empirical relevance to the early periods of the banking history during which a large number of banks issued their own notes and discounted others', such as the Free Banking Era in the United States (1838–63) and a period between 1750 and 1844 in England.¹⁵ Specifically, the baseline model derives a relationship between the loan interest rate r , the discount factor δ , the leverage ratio $\Lambda := D/G$, and the risks of default $1 - q$ of individual banks. First, $\frac{1+r}{\delta} = \widehat{R}$, a constant across banks, by the binding constraint (12). Therefore, the paper predicts that *across banks, the gross interest rate that a bank charges for loans (i.e., $1 + r$) is positively related to and solely determined by the discount factor of its notes δ , independent of the bank's other attributes*. Second, regarding the discount factor, from (10) and binding (12) it follows that

$$\delta = \left\{ \begin{array}{l} 1 \text{ if } \Lambda \leq \left(1 - \frac{A}{A\alpha}\widehat{R}\right)^{-1} \\ \frac{q+(1-q)\Lambda^{-1}}{1-(1-q)\frac{A}{A\alpha}\widehat{R}} \text{ if } \Lambda \geq \left(1 - \frac{A}{A\alpha}\widehat{R}\right)^{-1} \end{array} \right\}. \quad (13)$$

Therefore, *the discount factor of a bank's notes δ is inversely related to the bank's default risk $1 - q$ ¹⁶ and its leverage Λ (which is endogenous)*. The first part of this prediction is empirically confirmed by Gorton (1999), who studies the pricing of bank notes during the U.S. Free Banking Era.¹⁷

The applicability of the baseline model to modern banking is more limited. A fundamental assumption of the model, namely that banks face no constraints in making loans, is probably far from reality nowadays. In particular, due to this assumption, the baseline

¹⁵For the case of the United States, according to Gorton (1999), thousands of different banks' notes were in circulation, while for the case of England, Cameron et al. (1967) report that in year 1810 there were 783 note-issuing county banks in England.

¹⁶In the lower branch of (13) $\partial\delta/\partial q$ share the sign of $1 - \frac{A}{A\alpha}\widehat{R} - \Lambda^{-1}$, which is positive because in that branch $\Lambda \geq \left(1 - \frac{A}{A\alpha}\widehat{R}\right)^{-1}$, that is, $\Lambda^{-1} \leq 1 - \frac{A}{A\alpha}\widehat{R}$.

¹⁷Specifically, he found that possessing traits of low default risks, such as being a member of the Suffolk System or under a state-sponsored insurance, is negatively correlated with the *implied volatility* of the bank, of which the discount factor—or price, in his paper—is a decreasing function. Therefore, the low-risk traits are associated with a higher value of δ .

model predicts that the aggregate quantity of bank credit is independent of banks' wealth, G , whereas the aftermath of the recent financial crisis witnesses that the banking sector reduces credit issuance after suffering a severe loss. To give an account for this observation in particular and to examine the modern banking in general, we shall introduce a borrowing constraint to banks, as below.

4. Banks' Wealth Matters in the Presence of a Borrowing Constraint

In this section, we assume banks are subject to a borrowing constraint, that their equity value should never fall below $(1 - \mu)G$.¹⁸ The equity value is higher in the good state than it is in the bad state. Thus, we only need to consider the constraint in the bad state, which is $G + \underline{Y} - D \geq (1 - \mu)G$. With $\underline{Y} = \underline{A}(1 + r)/(\overline{A}\alpha) \times D$ by (8) and rearrangement, it becomes

$$D \cdot \left[1 - \frac{\underline{A}(1 + r)}{\overline{A}\alpha} \right] \leq \mu G. \quad (14)$$

The introduction of the constraint has two immediate implications. The first is that, because banks always maintain a positive equity value, they never default. Therefore, their notes are not discounted, that is, $\delta = 1$ and hence $1 + r = \widehat{R}$ hereafter. As a result, the borrowing constraint (14) becomes

$$D \cdot \left(1 - \frac{\underline{A}}{\overline{A}\alpha} \widehat{R} \right) \leq \mu G. \quad (15)$$

The other implication is that, unlike in the preceding case of least-fettered issuance, banks now are in Bertrand competition with *limited* capacities à la Kreps and Scheinkman (1983). With the borrowing constraint, the market equilibrium is as follows.

¹⁸This borrowing constraint can be due to several reasons. One reason is that, in modern times, bank default becomes very costly and banks' managers want to maintain solvency in any contingency, which is a special case of the above constraint with $\mu = 1$. Another reason, as in Gertler and Kiyotaki (2010) and Holmstrom and Tirole (1997), is that banks have a moral hazard issue: at $t = 1$, the owner of a bank—the banker—can abscond with a fraction of $1 - \mu$ of its stored wealth to a remote island. Hence, banks' equity value should never fall below $(1 - \mu)G$ in any contingency.

PROPOSITION 3.

(i) If $G \geq \frac{1}{\mu}G^*$, then in all equilibriums the real allocation is the same as in the case of least-fettered issuance: $\widehat{R} = R^{SB}$, the profit margin of bank lending is 0, and $L = L^{SB}$. If $G < \frac{1}{\mu}G^*$, there is a unique equilibrium in which the profit margin of bank lending is positive; and $\widehat{R} > R^{SB}$ and \widehat{R} is determined by G through

$$G = \frac{1}{\mu} \left(\frac{\overline{A}\alpha}{w^\alpha} \right)^{\frac{1}{1-\alpha}} \frac{1 - \frac{\overline{A}}{\overline{A}\alpha} \widehat{R}}{\widehat{R}^{\frac{1}{1-\alpha}}} \equiv G(\widehat{R}). \quad (16)$$

(ii) If G decreases from $\frac{1}{\mu}G^*$ to 0, the equilibrium interest rate, \widehat{R} , increases from R^{SB} to $\overline{A}\alpha/\underline{A}$ and the number of workers hired by entrepreneurs decreases from L^{SB} to $L^{SB} \times \left[\frac{\underline{A}}{q\overline{A}\alpha + (1-q)\underline{A}} \right]^{\frac{1}{1-\alpha}}$.

The borrowing constraint limits banks' capacity of issuance to an endogenous proportion of their wealth. If their wealth is high enough—that is, beyond $\frac{1}{\mu}G^*$ —then banks still possess a capacity large enough to annihilate the profit margin of issuance, which leads to the second-best allocation, as was in the case of least-fettered issuance. Hence arises result (i).

If $G < \frac{1}{\mu}G^*$, banks' wealth does not suffice to back an issuance of that size. As a result, issuance bears a positive profit margin. This positive profit margin drives all banks to issue as much as possible, until the borrowing constraint, (15), is binding. This clears the indeterminacy in the quantity of individual banks' issues and gives rise to a unique equilibrium. It also shows that, now, the quantity of money circulated is determined by the supply side—namely, the banking sector—rather than by the demand side, as was in the case of least-fettered issuance. Therefore, the lower the banks' wealth, the lower the total value of money that they supply and, as a result, the higher the interest rate of borrowing (\widehat{R}) and the fewer the workers that entrepreneurs hire.

The last comparative static, however, does not mean the aggregate output always decreases with G for $G < \frac{1}{\mu}G^*$ because of the wedge between the first-best and second-best allocations, namely $L^{FB} < L^{SB}$. Define

$$G^{FB} \equiv G(R^{FB}),$$

where function $G(\widehat{R})$ is defined in equation (16); with $R^{FB} = \bar{A}/A_e$,

$$G^{FB} = \frac{1}{\mu}(A_e\alpha - \underline{A})(A_e\alpha/w)^{\frac{\alpha}{1-\alpha}}.$$

Then G^{FB} is the level of banks' wealth at which the actual interest rate exactly takes the first-best value and therefore the aggregate output is maximized. $G^{FB} < \frac{1}{\mu}G^*$.¹⁹ By proposition 3, then, there are two types of efficiency:

- (i) $G < G^{FB}$. In this case, $\widehat{R} > R^{FB}$ and $L < L^{FB}$. That is, relative to the first-best allocation, banks under-lend, whereby the cost of bank credit is too high and the sector that depends on it—namely entrepreneurs—obtains inadequate resources, namely labor.
- (ii) $G > G^{FB}$. In this case, $\widehat{R} < R^{FB}$ and $L > L^{FB}$. That is, relative to the first-best allocation, banks over-lend, whereby the cost of bank credit is too low and the sector that depends on it obtains excessive resources.

The existence of the second type of inefficiency is due to assumption 3, which drives a wedge between the first-best and second-best allocations. In the absence of this wedge, R^{FB} would be equal to R^{SB} , hence $G(R^{FB})$ to $\frac{1}{\mu}G^*$, and the case of bank over-lending would not exist. This case of over-lending provides the paper with room to accommodate interest rate policy and capital adequacy regulation, as will be shown in section 6.

The first type of inefficiency is caused by inadequate creation of money by banks. Can the central bank enlarge the money supply through the banking system? It can, as we show in the next section, with a policy which, because it eases the constraints imposed by a dearth of money, is called the quantitative easing policy, or QE.

¹⁹That is because $R^{FB} > R^{SB}$, $\frac{1}{\mu}G^* = G(R^{SB})$ and function $G(R)$ is decreasing.

5. QE Reduces the Borrowing Constraint of Banks

The central bank (CB hereafter) in this economy is modeled as the unique entity that is able to costlessly produce another means of payment which has no intrinsic values; to fix the idea, let it be shells. Shells are different in nature to bank notes. The CB does not promise to pay a holder of shells with any corn, the real good. Shells, therefore, are purely nominal. By contrast, a bank commits to redeem its notes with the promised amounts of corn. Hence banks' notes are not nominal.

Although shells are fiat money and the model economy lasts for only two dates, shells can circulate in it, whereby the CB can conduct meaningful monetary policy, such as the following QE. At $t = 0$ the CB announces a facility whereby each of the N banks can borrow up to S units of shells (again, one unit is defined as N kg) at a small interest rate $\epsilon > 0$. A borrower bank is obliged to repay its debt to the CB at $t = 1$ either with shells or with corn, with a fixed rate of exchanging corn for shells—for example, 1 kg corn for 1 kg shells. This exchange rate is arbitrary and is made to confer a nominal value on shells, as in real life the Bank of England can print an arbitrary number, like £10, on a piece of paper, and then this piece of paper has a nominal value of £10 and can be used to buy real goods of this value (e.g., two packs of cherries at M&S). As such, we call 1 kg corn the *par value* of 1 kg shells, etc., although shells are nominal. The banks that have used the CB's facility thus have two types of money to lend to entrepreneurs. One is their own notes; the other is shells. We assume that no entrepreneurs borrow both types of money, in order to avoid the unnecessary complication of trying to determine which debt, of the two, is senior in the case of default.

The contract of borrowing a bank's notes is still represented by (E, r) , whereby the entrepreneur borrows the bank's notes of face value E at $t = 0$ and then owes it a debt of $E(1+r)$, which he repays at $t = 1$ with corn.

The contract of borrowing shells is similarly represented by (E_s, r_s) , whereby the entrepreneur borrows E_s kg shells at $t = 0$ and then owes the bank a debt of $E_s(1 + r_s)$, which he repays at $t = 1$ with either corn or shells, with 1 kg corn equivalent to 1 kg shells.

The timing of events at $t = 0$ is as follows. First, the CB chooses S and publicly announces the policy. Then, banks decide the quantity of shells to borrow from the CB, $Q \leq S$, and post offers of $(D, r; Q, r_s)$. Based on these offers, entrepreneurs decide which bank to go to and how much to borrow. Lastly, they use borrowed means of payment (bank notes or shells) to hire workers and start the production.

The timing of events at $t = 1$ is as follows. First, entrepreneurs produce corn. Second, the market for shells opens, on which the shell-borrowing entrepreneurs use corn to buy shells from workers. Let \bar{p}_1 (\underline{p}_1) denote the shell price in the good (bad) state. Third, entrepreneurs settle their debts to the lender banks, using corn and/or shells. Fourth, the market for shells *may* open the second time, on which banks having excessive shells sell to banks in shortage. Lastly, banks redeem their notes from workers and settle their debts to the CB. At this stage, the CB may end up holding a certain quantity of corn, which it will transfer to the agents of the economy before the consumption starts.

Note that there is no surprise or nominal rigidity associated with QE. All the private-sector agents (i.e., banks, entrepreneurs, and workers) made decisions after having observed the CB's move and they can freely adjust their decisions with it. Still, QE produces real effects in this economy in one equilibrium if $G < \frac{1}{\mu}G^*$.²⁰ For any $G < G^{FB}$, moreover, there exists a unique value of S at which this equilibrium attains the first-best allocation. This is the equilibrium that we discuss below.

In this equilibrium, at $t = 0$, workers believe that they will be able to use shells to buy corn at $t = 1$ and thus they accept a wage payment with shells; and at $t = 1$, indeed shells receive a positive real value, intuitively because they can be used to settle certain debts

²⁰ As shells are fiat money, there is another equilibrium in which workers disbelieve that shells will have real value at $t = 1$ and thus do not accept to be paid with shells at $t = 0$, which are thus not circulated. This type of equilibrium exists for fiat money in general. By contrast, it never exists for banks' notes: banks each have G units of corn, therefore their notes always have a positive real value at $t = 1$; indeed, if a bank issues $D \leq G$, workers know, without any equilibrium calculation, that its notes are worth par. This difference results from the difference in nature between shells and banks' notes stated at the beginning of this section.

that otherwise have to be settled with corn. To solve the equilibrium, we first examine what happens at $t = 1$ and then what happens at $t = 0$. Consider, first, the case in which S is below a threshold (which is given in proposition 4 below) and therefore $r_s > \epsilon$ in equilibrium. In this case, all the banks borrow to the full capacity of QE, namely, $Q = S$, and lend all the shells out to entrepreneurs. Therefore, at the beginning of $t = 1$, there are NS units of shells in workers' hands and the aggregate debt of the shell-borrowing entrepreneurs is $NS \times (1 + r_s)$ units.

In the good state, as no entrepreneurs default, the shell price $\bar{p}_1 = 1$. On the one hand, shells can never be worth more than their nominal values: if the entrepreneurs use 1 kg corn only to buy less than 1 kg shells, they will not buy them but instead use 1 kg corn to settle a debt of 1 kg shells to the banks. On the other hand, in the good state, it cannot be that $\bar{p}_1 < 1$ either. Otherwise, the shell-borrowing entrepreneurs would want only shells, not corn, to repay all their debts. Their aggregate demand for shells would thus be $NS(1 + r_s)$, as they do not default. This demand, with $r_s > 0$, is bigger than NS , the supply of shells, leaving the market uncleared.

As $\bar{p}_1 = 1$, the entrepreneurs are indifferent to repaying their banks with shells or with corn in the good state. As a result, some banks may end up with more than $S(1 + \epsilon)$ units of shells, some less. If so, the former wants to sell the excess to the latter and the second shell market opens. On this market, the equilibrium shell price, p_B , equals 1 also, following the same argument as above. On the one hand, it is impossible that $p_B > 1$. On the other hand, if $p_B < 1$, banks all want to use only shells, not corn, to settle their debts to the CB, and their aggregate demand would be $NS(1 + \epsilon)$, greater than the total supply, NS , thus leaving the market uncleared. This argument holds true, and hence $p_B = 1$ is the only market clearing price, so long as $\epsilon > 0$, however small. To simplify exposition, in what follows, we consider the limit case where $\epsilon \rightarrow 0$.²¹

²¹Note, however, that if $\epsilon = 0$, then there is indeterminacy: any price $p_B \in [0, 1]$ can clear this second shell market. If $\epsilon = 0$, the aggregate excess exactly equals the aggregate shortage. Then, the supply side is willing to sell all the excessive shells at any $p_B \geq 0$, while the demand side is willing to buy exactly the same amount at any $p_B \leq 1$.

In the bad state, the output is low. Hence $\underline{p}_1 < 1$, as will be verified later. Thus, the shell-borrowing entrepreneurs use only shells, no corn, to repay the banks. Furthermore, in this state entrepreneurs all default according to lemma 2. This means all their outputs are used to settle their debts to the banks. With these two facts put together, all the shell-borrowing entrepreneurs' output is used to buy shells. Denote by Y_s the total output of one bank's shell-borrowing entrepreneurs. Then the shell market clearing commands that $\underline{p}_1 \times NS = NY_s$, or

$$\underline{p}_1 = Y_s/S. \tag{17}$$

With a calculation similar to that leading to equation (8), we find $\underline{p}_1 = \underline{A}(1 + r_s)/(\overline{A}\alpha)$, which, as is shown below in proposition 4, is smaller than 1 indeed, as was intuitively argued. Moreover, from $\underline{p}_1 = Y_s/S$ it follows that the shell-borrowing entrepreneurs of one bank altogether buy $Y_s/\underline{p}_1 = S$ units of shells. Hence, each bank ends up with having S units of shell equally and the second shell market does not open in the bad state.

The value of 1 kg shells at $t = 0$, p_0 , is the mean of their values at $t = 1$:

$$p_0 = q \times 1 + (1 - q) \times \frac{\underline{A}(1 + r_s)}{\overline{A}\alpha}. \tag{18}$$

Workers thus accept a wage payment with w/p_0 kg shells at $t = 0$ in the equilibrium.

The effect of this policy is presented in the following proposition.

PROPOSITION 4. *Suppose $G < \frac{1}{\mu}G^*$.*

(i) *If QE's scale $S < \underline{S}(G) \equiv \frac{\overline{A}\alpha}{q(\overline{A}\alpha - \underline{A})}(\frac{1}{\mu}G^* - G)$, then there is a unique equilibrium in which all the S units of shells are in circulation. In this equilibrium, QE produces real effects: the actual interest rate, \widehat{R} , is determined by S through*

$$S = \left(\frac{\underline{A}}{w^\alpha}\right)^{\frac{1}{1-\alpha}} \cdot \frac{1 - (1 - q)\theta}{q} \left(\frac{1}{\theta^{\frac{1}{1-\alpha}}} - \frac{\mu G}{1 - \theta}\right) \equiv F(\theta; G),$$

with $\theta \equiv \frac{\underline{A}}{A\alpha}\widehat{R}$. (19)

Moreover, the interest rate of lending shells is $r_s = \frac{[q\bar{A}\alpha + (1-q)\underline{A}]\hat{R} - \bar{A}\alpha}{\bar{A}\alpha - (1-q)\underline{A}\hat{R}}$. It satisfies $r_s > 0$ and $\underline{A}(1+r_s)/(\bar{A}\alpha) < 1$, as was said.

(ii) If S increases, \hat{R} and r_s decrease and L increases. At $S = \underline{S}(G)$, $\hat{R} = R^{SB}$, $r_s = 0$ and $L = L^{SB}$.

QE works by increasing the real value of money that banks lend out at $t = 0$. That explains why, compared with the situation of its absence (namely with $S = 0$), the cost of borrowing money (\hat{R}) is reduced, whereby entrepreneurs hire more workers. And the larger the scale of QE (i.e., S), the greater are these effects (so long as S is below $\underline{S}(G)$).

The question is: how is QE able to make banks enlarge their lending scale, backed by the same amount of wealth, without breaking the borrowing constraint? To answer this question, consider a representative bank's balance sheet at $t = 1$, as shown in table 1.

Consider the bad state, in which the borrowing constraint is binding. The constraint commands that $G + \underline{Y} + Y_s - D - \underline{p}_1 S \geq (1 - \mu)G$. This inequality, with $\underline{p}_1 = Y_s/S$ by (17), is equivalent to $G + \underline{Y} - D \geq (1 - \mu)G$, which leads to inequality (14), the same borrowing constraint in the absence of QE. Therefore, the injection of the fiat money with QE does not reduce the capacity of private issuance at all. This is due to $\underline{p}_1 = Y_s/S < p_0$,²² which means that in the bad state the shell price goes down, so that the real value of a bank's liability to the CB, $\underline{p}_1 S$, falls enough to be covered by the value of the assets, Y_s . The decrease in the real value of shells means inflation. Put differently, QE enables banks to lend more without breaking the borrowing constraint because it induces inflation in the bad state, which lowers the real value of banks' liability to the CB. The inflation arises because shells are fiat money. The CB never commits to redeem shells with a real good. Their value, therefore, freely adjusts with the state of the economy. By contrast, banks commit to redeem their notes (i.e., their promise to pay) at

²² $\underline{p}_1 < p_0 \Leftrightarrow \underline{p}_1 < q \times 1 + (1 - q) \times \underline{p}_1 \Leftrightarrow \underline{p}_1 < 1$, which holds true by the proposition.

Table 1. The Balance Sheet (in real value) of the Representative Bank with QE

Assets	Liabilities
Corn Stored (G)	Equity
Loans in Notes (Y)	Liability to the Note Holders (D)
Loans in Shells (Y_s)	Liability to the CB ($p_1 S$)

a pre-specified, non-contingent, value.²³ This difference in nature between banks' notes and shells, therefore, drives the functioning of QE.

Remark: An abstraction of the model helps elucidate this mechanism of QE. That is the denomination of banks' liabilities with corn, the real good. In reality, however, typically they are denominated with the fiat money currency of the economy. Incorporating this feature would not necessarily invalidate the working of the mechanism. First, the aforementioned difference in nature between bank-created money and currency is still there. Although banks now use the currency to redeem their liability, to obtain the currency they need to expend real resources. Therefore, their liability still bears real obligations of repayment. In contrast, fiat money is not redeemable and bears no such obligations. Second, it is true that if banks' liability is denominated with the currency, its real value falls under the negative shocks, which improves banks' conditions. However, a credit crunch still happens. So long as they face a binding borrowing constraint, they will still reduce lending if their wealth is profusely eaten off, as is evident in the aftermath of the 2008–09 crisis. Third, in this circumstance, if QE is able to raise banks' lending capacities, it has to ease their borrowing constraint. As it does not give real resources to banks, to do so it has to further reduce the real value of banks'

²³In the model economy, the value of these issues at $t = 1$ could be made contingent on the state in two ways. One is to let the specified value be contingent on the state, which is disallowed by assumption 2. The other is default, which is disallowed because of the borrowing constraint, (14). While default is disallowed altogether because of the particular form that the constraint takes, in general so long as any borrowing constraint restricts banks' capacity, there is only a limited extent to which the value of their liabilities can adjust to the real economic conditions.

liability, and this can only be done via inflation. Fourth, this reduction can be done by the CB with QE, but not by banks themselves, exactly because of the above said difference in nature between the two types of money.

At the core is the point that this difference gives the CB leeway to act that banks do not have. This point can be seen even more clearly in a modified version of the setting where the CB implements QE by directly lending shells to entrepreneurs (rather than through banks) at interest rate r_s , which would result in the same prices of shells and the same allocation as in the original setting. Obviously, in this setting the CB can lend out the currency, whereas banks cannot lend more of their liabilities, because shells are not redeemable but banks' liabilities are. Moreover, if banks' liabilities were now denominated with the currency, this policy would still work. Actually it would work better because it would ease banks' borrowing constraint by inducing inflation, which enables them to enlarge the lending capacity.

Returning to the model, while QE induces inflation in the bad state, it causes deflation in the good state: $\bar{p}_1 = 1 > p_0$. This result might partly explain the lingering deflation pressure or the low inflation observed during or after the implementation of QE in several economies, such as the United Kingdom and the euro zone. This phenomenon would be a puzzle if considered from the point of view of the quantitative theory of money, given that QE enormously expands the monetary base.

Thus far we have considered the case in which $S < \underline{S}(G)$ and therefore $r_s > 0$. Now we consider what happens if QE is in a bigger scale. At $S = \underline{S}(G)$, $\hat{R} = R^{SB}$ and $r_s = 0$. By proposition 1(iii), \hat{R} can never fall below R^{SB} , otherwise banks would stop lending altogether, not a case in equilibrium. Therefore, if the scale of QE is larger than $\underline{S}(G)$, then \hat{R} stays at R^{SB} , r_s stays at 0, and the excessive shells beyond the threshold are not put into circulation; indeed, with $r_s = 0$, banks are indifferent regarding the quantity of shells to borrow and lend. This result partly offers an explanation for the phenomenon that, in certain economies such as the United States and the euro zone, a substantial fraction of monetary bases created with QE is not lent out but stays in banks' reserve accounts.

Now consider the optimal scale of QE. The first-best value of the interest rate is $\widehat{R} = R^{FB}$. In the circumstance where $G \geq G^{FB}$ and hence $\widehat{R} \leq R^{FB}$ at $S = 0$, any $S > 0$ only makes \widehat{R} still smaller and efficiency even lower. Therefore, the optimal $S = 0$, namely, the CB should not implement QE at all in the circumstance of banks over-lending. However, in the circumstance where $G < G^{FB}$ and hence $\widehat{R} > R^{FB}$ at $S = 0$, there is a unique value of S at which QE drags down \widehat{R} to R^{FB} . This is summarized below.

PROPOSITION 5.

- (i) If $G \geq G^{FB}$, the optimal scale of QE is $S = 0$, that is, no QE should be implemented.
- (ii) If $G < G^{FB}$, the optimal $S = F(\frac{A}{A\alpha} \cdot R^{FB}; G) > 0$, where function $F(\cdot)$ is given by (19). At this scale QE attains the first-best allocation (i.e., $\widehat{R} = R^{FB}$). Moreover, with optimal QE, banks earn profit from lending shells (i.e., $r_s > 0$), but their overall profit is reduced compared with the case without QE if $G \geq \underline{G}$ for some $\underline{G} < G^{FB}$.

According to the last result, although with QE banks receive free funding from the CB and lend it at a positive interest rate, surprisingly they lose from the policy, which thus does not subsidize them. That is because to banks, besides this positive effect on the scale, QE induces a negative, general equilibrium effect on the profit margin by enlarging the lending capacities of all banks. As a result, banks obtain a reduced profit from lending their notes. As $\partial F / \partial G < 0$, if banks' wealth is above a threshold (i.e., \underline{G}), then the scale of optimal QE is small enough. As a result, the positive effect from the enlargement of the lending scale is small and dominated by the negative effect from the reduction of the profit margin. However, even in this case of banks all losing from the free funding of the CB, given $r_s > 0$, individual banks still strictly prefer taking it rather than abstaining from it—thus QE can be conducted on a voluntary basis. The reason is that a single bank takes the profit margin of lending as given and neglects the effect on it of enlarging its own capacity.

In the circumstance where $G > G^{FB}$ and banks over-lend, we have seen that QE does not help. What the CB can do is discussed in the next section.

6. Interest Rate Policy and Capital Adequacy Regulation to Curb Bank Lending

If $G > G^{FB}$, then banks over-lend, making the actual interest rate too low, and entrepreneurs hire too many workers relative to the first-best allocation. In this circumstance, it is usually expected that the CB can help by setting a high policy interest rate. This idea is explored in this section, where we show that interest rate policy produces real effects and is able to attain the first-best allocation if and only if there is a nominal rigidity. In its absence, we show, capital adequacy regulation is always a remedy.

To explain the meaning of nominal rigidity in this paper, we define nominal wage as the total face value of banks' notes that workers receive as the wage payment. In the absence of any intervention from the CB, the nominal wage is w because banks do not default (due to the borrowing constraint (14)). The nominal rigidity in this paper is defined as the friction that keeps the nominal wage of workers staying at w and unchanged with the CB's policy.

6.1 Interest Rate Policy Works If and Only If the Nominal Rigidity Is Present

In the two-date economy of this paper, the CB sets the policy rate r_p by offering to workers a savings account which takes in the deposit of banks' notes at $t = 0$ and pays out with shells at $t = 1$. Specifically, if the CB receives a deposit of some banks' notes of overall face value F kg corn at $t = 0$, it issues to the depositor $F(1 + r_p)$ kg shells at $t = 1$. Moreover, by taking in the notes, the CB becomes a creditor to the issuer banks and charges these banks an interest rate of $1 + r_p + \varepsilon$ for some $\varepsilon > 0$. Thus it obliges these banks altogether to pay back $F(1 + r_p + \varepsilon)$ at $t = 1$, either with corn or with shells, counting 1 kg corn equivalent to 1 kg shells. There is one equilibrium in which this policy is meaningful.²⁴ In this equilibrium *if* the aggregate face value of notes deposited with the CB at $t = 0$ is F and consequently $F(1 + r_p)$ kg shells are created at $t = 1$, then these

²⁴As shells are fiat money, there is an equilibrium in which no one believes that shells have real value at $t = 1$ and thus no one deposits any bank notes with the CB at $t = 0$, which makes the policy rate meaningless.

shells are priced at par (i.e., 1 kg shells worth 1 kg corn) at $t = 1$. On the one hand, they can never be priced above par. On the other hand, shells are not priced below par either. Otherwise, the banks indebted to the CB would want to use only shells to settle all their debts. Thus, their aggregate demand for shells would be $F(1+r_p+\varepsilon)$ kg, but only $F(1+r_p)$ kg shells are created, insufficient to meet the demand. The deposit of bank notes with the CB and the creation of shells, however, will not actually happen in equilibrium. Banks will pay their note holders the same interest rate of r_p (or even a little more) to stop them from depositing the notes with the CB and avoid the payment of the even higher interest rate of $r_p + \varepsilon$. As a result of the policy rate set at r_p , therefore, a bank's note of face value 1 issued at $t = 0$ is worth $1 + r_p$ at $t = 1$, and banks that issue notes of aggregate face value D at $t = 0$ are in a liability of $D(1 + r_p)$ at $t = 1$.

Having explained how the interest rate policy is conducted in this economy, we state its effects in the following proposition.

PROPOSITION 6. *Suppose $G > G^{FB}$.*

- (i) *If workers' nominal wage adjusts with the policy rate, the interest rate policy produces no real effects but deflates the nominal wage to $w/(1 + r_p)$.*
- (ii) *If workers' nominal wage stays at w invariant to the CB's policy rate, the equilibrium interest rate of bank lending, \widehat{R} , increases with the policy rate r_p . At a unique value of r_p , $\widehat{R} = R^{FB}$ and the first-best allocation is attained. With the optimal interest rate policy, banks obtain zero profit if $G \geq G_s$, where $G_s \equiv \frac{1}{\mu}[(q\overline{A}\alpha + (1 - q)\underline{A}) - \underline{A}](A_e\alpha/w)^{\frac{\alpha}{1-\alpha}}$ and satisfies $G^{FB} < G_s < \frac{1}{\mu}G^*$.*

Intuitively, in both the regime of flexible nominal wage and that of sticky nominal wage, a higher policy rate, by increasing the cost of deposit, induces the bank to set a higher lending rate in the hope of shifting the increased cost of deposit to entrepreneurs: r increases with r_p in both regimes, as shown in the proof. As a result, in both regimes, a higher policy rate reduces the quantity of bank credit that entrepreneurs borrow. In the regime of flexible wage, this reduction

produces no effect on the number of workers they employ because it is exactly offset by the decrease in the nominal wage that workers accept. In the regime of sticky wage, by contrast, the nominal wage stays the same and this reduction in bank credit forces entrepreneurs to hire fewer workers. With a proper policy rate, the number of workers they hire is brought down to the first-best level, that is, the first-best allocation is attained. The decrease in entrepreneurs' demand for bank credit subjects banks to stronger competition. As a result, their profit margin of lending is nullified if their lending capacity is not too small, that is, if their wealth is not too low—i.e., no less than G_s .

The interest rate policy, therefore, works to improve efficiency over the market equilibrium if and only if the nominal rigidity is present. However, regulation that sets a lower bound on banks' capital adequacy ratio always works to curb the over-lending by banks (if the CB has the authority for such regulation), as is shown below.

6.2 Capital Adequacy Regulation Always Works

The aggregate face value of money that is needed for entrepreneurs to hire the first-best number of workers is $wL^{FB} := D^{FB}$. Suppose the CB imposes the constraint that banks cannot issue more than D^{FB}/G times their wealth, G . Then, this constraint is binding in the competitive equilibrium if $G > G^{FB}$ because, in its absence, banks issue more than D^{FB} (as $\widehat{R} < R^{FB}$) by proposition 3, and thus the constraint is violated. The constraint, therefore, is binding, so banks issue D^{FB} and the first-best allocation is attained. One way to implement this constraint is to impose at $t = 0$ a lower bound on banks' capital adequacy ratio, which is defined as the equity-to-asset ratio in market value, that is, $(G + Y - D)/(G + Y)$, where Y is the value of banks' loans at $t = 0$, that is, $Y = q \times \overline{Y} + (1 - q) \times \underline{Y}$. Define

$$c^{FB} := \frac{A_e \alpha G + (1 - q) \underline{A} (1 - \alpha) D^{FB}}{A_e \alpha G + [q \overline{A} \alpha + (1 - q) \underline{A}] D^{FB}},$$

which is the capital adequacy ratio if banks lend the first-best quantity of money (as is shown in the proof of the following proposition).

PROPOSITION 7. *Suppose $G > G^{FB}$. If the CB imposes regulation that restricts banks' capital adequacy ratio from being smaller than c^{FB} , then banks issue the first-best quantity of money and the first-best allocation is attained.*

The intuition for why the regulation works is simple: the money that banks create is their liability. Therefore, the scale of its issuance is subject to the capital adequacy regulation. If the regulation is tight, then banks are restricted from over-lending.

7. Conclusion

Banks are special because their liability, especially that in the form of demand deposit, is widely accepted as a means of payment. Often, real sectors need to borrow banks' liability (so-called *money* or *credit*) to obtain resources that they want. This interaction between the real sectors and the banking sector on competitive markets, as well as the central bank's remedy to market inefficiency, is the focus of this paper. It underlines the importance of banks' aggregate wealth. Depending on the size of this wealth, there are two types of inefficiency.

If banks' wealth is below a threshold, then they issue too little money, with symptoms that the interest rate of bank credit is too high and the sectors that depend on it obtain inadequate resources. In this circumstance, the central bank can improve efficiency by lending to all banks its fiat money at zero interest rate. This policy, which bears the flavor of QE, works because of a difference in nature between bank-created money and fiat money. The latter is not redeemable, whereas the former bears real obligations of redemption. Furthermore, the policy induces deflation under the positive productivity shock. Lastly, while the policy gives to banks free funding, which they lend out at a positive interest rate, it does not subsidize them unless their wealth is very low.

If banks' wealth is above the threshold, on the other hand, banks lend out too much money, with symptoms that the interest rate of bank credit is too low and the resources are skewed to the sectors that highly depend on it. To curb over-lending by banks, the central bank may set a high policy rate, which, however, works if and only

if the nominal wage cannot freely adjust with the policy. By contrast, the regulation that sets a tight capital adequacy ratio always works because the money that banks create is their liability, and its quantity is thus subject to the regulation.

Appendix: The Proofs

Proof of Lemma 1

In equilibrium, only one promised wage, denoted by F , prevails on the market, as will be shown. Competitive equilibrium is thus defined as a profile of (F, L) , such that

- (i) given that F prevails on the market, the optimal demand for labor of each entrepreneur is L ;
- (ii) given that each entrepreneur demands L workers, F clears the labor market.

The two conditions are elaborated as follows.

For (i), given F , a representative entrepreneur's decision problem on labor demand is

$$\max_L q(\overline{A}L^\alpha - FL) + (1 - q)\max(\underline{A}L^\alpha - FL, 0),$$

where the "max" term appears because the entrepreneur might default in the bad state. That is indeed the case at the optimum. Otherwise, the entrepreneur's problem is

$$\max_L q(\overline{A}L^\alpha - FL) + (1 - q)(\underline{A}L^\alpha - FL).$$

The solution is $L = (\frac{A_e\alpha}{F})^{\frac{1}{1-\alpha}}$. Then in the bad state his output is $\underline{A}((\frac{A_e\alpha}{F})^{\frac{1}{1-\alpha}})^\alpha$, which is smaller than $F \cdot (\frac{A_e\alpha}{F})^{\frac{1}{1-\alpha}}$, the wage obligation, because $\underline{A} < A_e\alpha$ as assumed in (1). Hence, he defaults in the bad state, contradictory to what was supposed.

Defaulting in the bad state, entrepreneurs choose L to maximize the profit in the good state, $\overline{A}L^\alpha - FL$. Therefore, given F , the labor demand is

$$L = \left(\frac{\overline{A}\alpha}{F}\right)^{\frac{1}{1-\alpha}}. \quad (20)$$

For (ii), as there are a lot more workers than entrepreneurs can hire, the labor market is cleared by an expected wage income of w , the output of workers in autarky. In the good state, the workers hired get the promised wage, F . In the bad state, entrepreneurs default and all the output goes to the workers, of which each obtains $\frac{\underline{A}L^\alpha}{L} = \underline{A}L^{\alpha-1}$. The labor market clears if

$$qF + (1 - q)\underline{A}L^{\alpha-1} = w. \tag{21}$$

Equations (20) and (21) together pin down L^{SB} as given in the lemma.

Now, show that only one F prevails on the market. If an entrepreneur posts F , then by (20) he hires $L = (\frac{\bar{A}\alpha}{F})^{\frac{1}{1-\alpha}}$ workers, whose wage income is F in the good state and $\underline{A}L^{\alpha-1} = \frac{\underline{A}}{\bar{A}\alpha}F$ in the bad state. Both increase with F . Therefore, workers go only to entrepreneurs who post the highest F , and in competitive equilibrium, only one F prevails.

Q.E.D.

Proof of Lemma 2

Suppose, otherwise, an entrepreneur does not default in the bad state. Then, his problem is

$$\max_E q(\bar{A}L^\alpha - E(1 + r)) + (1 - q)(\underline{A}L^\alpha - E(1 + r)) \text{ s.t. (3)}.$$

From the constraint, $E = wL/\delta$. Substitute it into the objective and let $\gamma \equiv w(1 + r)/\delta$. Then, the problem becomes

$$\max_L A_e L^\alpha - \gamma L.$$

The solution is $L = (A_e\alpha/\gamma)^{\frac{1}{1-\alpha}}$. At this scale, the entrepreneur will default in the bad state: $\underline{A}L^\alpha < E(1 + r)|_{E=wL/\delta; \gamma \equiv w(1+r)/\delta} \Leftrightarrow \underline{A}L^\alpha < \gamma L \Leftrightarrow \underline{A}L^{\alpha-1} < \gamma|_{L=(A_e\alpha/\gamma)^{\frac{1}{1-\alpha}}} \Leftrightarrow \underline{A} < A_e\alpha$, which is assumed in (1)—hence a contraction to what was supposed.

Q.E.D.

Proof of Lemma 3

As the entrepreneurs all default in the bad state and each hands over his whole output, y , to the bank, the value of the bank's loans in the bad state, \underline{Y} , equals y times the number of entrepreneurs that the bank finances, D/E . With $y = \underline{A}L^\alpha$, and E and L given by (4) and (5),

$$\frac{y}{E} = \frac{\underline{A}(1+r)}{\bar{A}\alpha}. \tag{22}$$

Then, $\underline{Y} = y \cdot D/E = D \cdot y/E = \underline{A}(1+r)/\bar{A}\alpha \cdot D$, that is, (8).
 Q.E.D.

Proof of Lemma 4

$\Pi(D, r) = E_{\tilde{Y}} \max(G + \tilde{Y} - D, 0) - G = E_{\tilde{Y}} \max(\tilde{Y} - D, -G) = E_{\tilde{Y}} \left[\tilde{Y} + \max(-D, -G - \tilde{Y}) \right] = E_{\tilde{Y}} \left[\tilde{Y} - \min(D, G + \tilde{Y}) \right] = E_{\tilde{Y}} \left[\tilde{Y} - D \times \min(1, \frac{G+\tilde{Y}}{D}) \right] = D \times E_{\tilde{Y}} \left[\frac{\tilde{Y}}{D} - \min(1, \frac{G+\tilde{Y}}{D}) \right]$, which, with $\delta = \min(1, \frac{G+\tilde{Y}}{D})$, equals $D \times E_{\tilde{Y}} \left[\frac{\tilde{Y}}{D} - \delta \right] = D \times \left[q \frac{\tilde{Y}}{D} + (1-q) \frac{Y}{D} - \delta \right]$, which, as $\bar{Y} = D(1+r)$ and $\underline{Y} = D \times \frac{\underline{A}}{\bar{A}\alpha} (1+r)$, equals $D \times \left[q(1+r) + (1-q) \frac{\underline{A}}{\bar{A}\alpha} (1+r) - \delta \right] = D \times \left[\frac{1+r}{R^{SB}} - \delta \right]$.
 Q.E.D.

Proof of Proposition 2

Proof of proposition 2(i) has been shown in the main text. For proposition 2(ii), prove the “if” part by reduction to absurdity. Suppose that in one equilibrium, no banks default; namely, (9) is honored for all banks. Then, for all banks $\delta = 1$ and $1+r = \hat{R}|_{\text{result } (i)} = R^{SB}$. By (9), each bank issues

$$D \leq G / \left(1 - \frac{\underline{A}}{\bar{A}\alpha} R^{SB} \right) = \frac{G[q\bar{A}\alpha + (1-q)\underline{A}]}{q(\bar{A}\alpha - \underline{A})}.$$

With $\delta = 1$ and $1+r = R^{SB}$, by (4) the demand by entrepreneurs is $E = (q\bar{A}\alpha + (1-q)\underline{A})^{\frac{1}{1-\alpha}} w^{\frac{\alpha}{1-\alpha}}$. If $G < G^*$, then $D < E$, that is, the

supply is below the demand—thus not in equilibrium. To prove the “only if” part of proposition 2(ii), it suffices to show that if $G \geq G^*$, no banks default in the symmetric equilibrium, which is constructed as follows. Banks all choose $r = R^{SB} - 1$ and D to satisfy entrepreneurs' demand for notes $E = (q\bar{A}\alpha + (1 - q)\underline{A})^{\frac{1}{1-\alpha}} w^{\frac{-\alpha}{1-\alpha}}$. With this value of (D, r) , it is straightforward to check that if $G \geq G^*$, the no-default condition, (9), is honored and hence no banks default in the bad state. Q.E.D.

Proof of Proposition 3

Note that, as there is no default, $\delta = 1$ and $1 + r = R$. In any equilibrium, the actual interest rate that all banks offer equals the market clearing rate, that is, $R = \hat{R}$. As $1 + r = \hat{R}$ and $\delta = 1$, by (4), the demand by entrepreneurs is $E = [\bar{A}\alpha/(\hat{R}w^\alpha)]^{\frac{1}{1-\alpha}} := E(\hat{R})$, which decreases with \hat{R} . Due to the borrowing constraint, (14), and with $R = \hat{R}$, the quantity of notes supply satisfies

$$D \leq \mu G \left(1 - \frac{A}{A\alpha} \hat{R}\right)^{-1} := \bar{D}(\hat{R}; G).$$

Moreover, if $\hat{R} > R^{SB}$, the profit margin of lending is positive, banks want to lend as much as they can, and therefore the constraint is binding. It follows that the supply function is

$$D(\hat{R}; G) = \left\{ \begin{array}{l} 0 \text{ if } \hat{R} < R^{SB} \\ [0, \bar{D}(\hat{R}; G)] \text{ if } \hat{R} = R^{SB} \\ \bar{D}(\hat{R}; G) \text{ if } \hat{R} > R^{SB} \end{array} \right\}.$$

The equilibrium value of \hat{R} is determined by $E(\hat{R}) = D(\hat{R}; G)$.

For proposition 3(i), if $G \geq \frac{1}{\mu}G^*$, then $D(\hat{R}; G) > E(\hat{R})$ for $\hat{R} > R^{SB}$ and $D(\hat{R}; G) < E(\hat{R})$ for $\hat{R} < R^{SB}$. Therefore, in equilibrium $\hat{R} = R^{SB}$. Hence $L = L^{SB}$ —the second-best allocation is attained—and the profit margin of lending is 0, both of which hold in the case of least-fettered issuance.

If $G < \frac{1}{\mu}G^*$, $D(\widehat{R}; G) < E(\widehat{R})$ for $\widehat{R} \leq R^{SB}$. Therefore, in equilibrium $\widehat{R} > R^{SB}$ and is determined by $E(\widehat{R}) = \overline{D}(\widehat{R}; G)$ or, equivalently, equation (16), replicated below:

$$G = G(\widehat{R}) \equiv \frac{1}{\mu} \left(\frac{\overline{A}\alpha}{w^\alpha} \right)^{\frac{1}{1-\alpha}} \cdot \frac{1 - \frac{A}{\overline{A}\alpha} \widehat{R}}{\widehat{R}^{\frac{1}{1-\alpha}}}. \tag{23}$$

It is straightforward to show for this function $G(\cdot)$ that $G' < 0$, $G(R^{SB}) = \frac{1}{\mu}G^*$ and $G(\overline{A}\alpha/\underline{A}) = 0$.

Furthermore, as all the banks issue \overline{D} , there is no indeterminacy in the scale of issuance by individual banks and the equilibrium uniquely exists.

To prove proposition 3(ii), let $\widehat{R}(G)$ be the inverse function of $G(\widehat{R})$. Then, in the equilibrium $\widehat{R} = \widehat{R}(G)$ if $G < \frac{1}{\mu}G^*$. As $G(\widehat{R})$ is decreasing, so is $\widehat{R}(G)$. Moreover, $\widehat{R}(\frac{1}{\mu}G^*) = R^{SB}$ and $\widehat{R}(0) = \overline{A}\alpha/\underline{A}$ because $G(R^{SB}) = \frac{1}{\mu}G^*$ and $G(\overline{A}\alpha/\underline{A}) = 0$. With $R = \widehat{R}$, the number of workers hired in equilibrium, by (5), is $L = (\overline{A}\alpha/w\widehat{R})^{\frac{1}{1-\alpha}}$, which decreases with \widehat{R} . Thus, L increases with G . Moreover, $L = L^{SB}$ at $\widehat{R} = R^{SB}$ which holds if $G = \frac{1}{\mu}G^*$. At the other end, if $G = 0$, then $\widehat{R} = \overline{A}\alpha/\underline{A}$. As L is in proportion to $(1/\widehat{R})^{\frac{1}{1-\alpha}}$, we have $L = L^{SB} \times (R^{SB}/\widehat{R})^{\frac{1}{1-\alpha}} = L^{SB} \times [\frac{R^{SB}}{\overline{A}\alpha/\underline{A}}]^{\frac{1}{1-\alpha}} = L^{SB} \times [\frac{A}{q\overline{A}\alpha + (1-q)\underline{A}}]^{\frac{1}{1-\alpha}}$.

Q.E.D.

Proof of Proposition 4

For proposition 4(i), to characterize the equilibrium, observe that if QE’s scale S is small enough—the threshold for which will be found later—then $\widehat{R} > R^{SB}$ still holds and $r_s > 0$. That is, lending bears a positive profit margin. It has two implications: (a) both notes and shells are lent out; and (b) they are lent in the maximum quantity, that is, banks’ borrowing constraint (15) is binding and all the S units of shells are in circulation.

From implication (a) it follows that the actual interest rates of borrowing the two types of money are equalized, and equal to \widehat{R} , the

market rate. Following the discussion in subsection 3.1, the actual interest rate of borrowing shells is $(1+r_s)/p_0$, while the actual interest rate of borrowing banks' notes is $1+r$ (since $\delta = 1$ as banks do not default). Therefore,

$$\frac{1+r_s}{p_0} = 1+r = \widehat{R}. \tag{24}$$

From implication (b), the quantity of notes issued is $D = \mu G / (1 - \frac{A}{A\alpha} \widehat{R})$. Given that they are not discounted (i.e., $\delta = 1$) and the ex ante value of shells per unit is p_0 , the aggregate value of means of payment supplied is $p_0 S + D$, which, when the market clears, equals the wage payment that entrepreneurs demand to hire workers:

$$wL = p_0 S + \frac{\mu G}{1 - \frac{A}{A\alpha} \widehat{R}}. \tag{25}$$

By (5), the number of workers they hire is

$$L = \left(\frac{\overline{A\alpha}}{w} \right)^{\frac{1}{1-\alpha}} \widehat{R}^{\frac{-1}{1-\alpha}}. \tag{26}$$

These four equations (note that (24) has two) together with equation (18) (which settles p_0), as shown below, determine a unique profile of $(p_0, r_s, r, \widehat{R}, L)$ in equilibrium—thus, the equilibrium in which shells circulate exists uniquely. Moving on to show that, we first derive equations (19). By (24), $1+r_s = \widehat{R}p_0$. Substituting it into (18) and rearranging, we have

$$p_0 = \frac{q}{1 - (1-q)\frac{A}{A\alpha} \cdot \widehat{R}}. \tag{27}$$

Substitute it and (26) into (25), rearrange, let $\theta \equiv \frac{A}{A\alpha} \cdot \widehat{R}$, and we come to (19):

$$S = \left(\frac{A}{w^\alpha} \right)^{\frac{1}{1-\alpha}} \frac{1 - (1-q)\theta}{q} \left(\frac{1}{\theta^{\frac{1}{1-\alpha}}} - \frac{\mu G}{1-\theta} \right) \equiv F(\theta). \tag{28}$$

It is straightforward to verify that (a) $F'(\theta) < 0$; (b) equation $F(\theta) = 0$ is equivalent to (16). Thus, at $S = 0$, $\widehat{R} = \widehat{R}(G)$,

where $\widehat{R}(G)$ is the interest rate determined by (16) in proposition 3, namely the actual interest rate without QE; and (c) at $\widehat{R} = R^{SB}$, $F(\frac{A}{A\alpha}R^{SB}) = \underline{S}(G)$. It follows that for any $S < \underline{S}(G)$, equation (28) determines a unique $\widehat{R} \in (R^{SB}, \widehat{R}(G)]$; hence $\underline{S}(G)$ is the threshold that was to be found at the beginning of the proof. After \widehat{R} is found, it uniquely determines r , L , and p_0 , respectively, through equations (24), (26), and (27). It also uniquely determines r_s by putting equations (24) and (27) together, which leads to

$$r_s = \frac{[q\bar{A}\alpha + (1 - q)\underline{A}]\widehat{R} - \bar{A}\alpha}{\bar{A}\alpha - (1 - q)\underline{A}\widehat{R}}. \tag{29}$$

Therefore, any $S \in [0, \underline{S}(G))$ pins down a unique equilibrium profile of $(p_0, r_s, r, \widehat{R}, L)$.

By (29), $r_s > 0 \Leftrightarrow \widehat{R} > R^{SB} \Leftrightarrow S < \underline{S}(G)$. Moreover, $\frac{\underline{A}(1+r_s)}{\bar{A}\alpha} < 1 \Leftrightarrow 1+r_s < \bar{A}\alpha/\underline{A}|_{(29)} \Leftrightarrow q\widehat{R}/[1 - (1-q)\underline{A}/(\bar{A}\alpha) \cdot \widehat{R}] < \bar{A}\alpha/\underline{A} \Leftrightarrow \widehat{R} < \bar{A}\alpha/\underline{A}|_{\widehat{R} < \widehat{R}(G)} \Leftrightarrow \widehat{R}(G) < \bar{A}\alpha/\underline{A}$, which is affirmed by proposition 3(ii).

For proposition 4(ii), as $F'(\theta) < 0$, we have θ , and thus \widehat{R} , decrease with S . By property (c) of function $F(\cdot)$ above, $\widehat{R} = R^{SB}$ at $S = \underline{S}(G)$. By (29), r_s increases with \widehat{R} and $r_s = 0$ at $\widehat{R} = R^{SB}$. Therefore, r_s decreases with S and equals 0 at $S = \underline{S}(G)$.

Q.E.D.

Proof of Proposition 5

Result (i) is proved in the main text. As for (ii), as $G < G^{FB}$, $\widehat{R}(G) > R^{FB} > R^{SB}$. Note that $\widehat{R} = \widehat{R}(G)$ at $S = 0$ and $\widehat{R} = R^{SB}$ at $S = \underline{S}(G)$. Therefore, there is a unique S between 0 and $\underline{S}(G)$ at which $\widehat{R} = R^{FB}$ and this S equals $F(\frac{A}{A\alpha} \cdot R^{FB})$ by (19). As for the profit of banks, in the unique equilibrium, each bank serves N entrepreneurs and the profit obtained from one of them, $\widehat{\pi}$, is the difference of the social value of his project minus his profit from it; that is, $\widehat{\pi} = A_e L^\alpha - wL - V$. With L and V given by (5) and (6) and $R = \widehat{R}$,

$$\widehat{\pi} = \left(\frac{\bar{A}\alpha}{w}\right)^{\frac{\alpha}{1-\alpha}} (q\bar{A}\alpha + (1 - q)\underline{A}) \cdot \widehat{R}^{\frac{-1}{1-\alpha}} (\widehat{R} - R^{SB}).$$

Note that $\hat{\pi}$ increases with \hat{R} for $\hat{R} \in [R^{SB}, \frac{1}{\alpha}R^{SB}]$. Define $\underline{G} \equiv G(\frac{1}{\alpha}R^{SB})$. Then, if $G \geq \underline{G}$, we have $\hat{R} \leq \hat{R}(G) \leq \hat{R}(\underline{G}) = \frac{1}{\alpha}R^{SB}$ for any $S \geq 0$. Therefore, QE of any $S > 0$ —in particular, the optimal one—lowers $\hat{\pi}$ and hence reduces banks' profit by decreasing \hat{R} .

Q.E.D.

Proof of Proposition 6

- (i) The equilibrium is found in a way parallel to that of proposition 3: first find the demand for notes given \hat{R} , then the supply, and finally the equilibrium value of \hat{R} . Given \hat{R} , entrepreneurs hire $L = (\frac{\bar{A}\alpha}{w})^{\frac{1}{1-\alpha}} \hat{R}^{\frac{-1}{1-\alpha}}$ workers by (5). With flexible nominal wage, workers accept a nominal wage of $w/(1+r_p)$. Therefore, the demand for notes is

$$E^* (\hat{R}) = w \left(\frac{\bar{A}\alpha}{w} \right)^{\frac{1}{1-\alpha}} \frac{\hat{R}^{\frac{-1}{1-\alpha}}}{1+r_p}.$$

Now consider the supply side. By borrowing face value E , an entrepreneur hires $E(1+r_p)/w$ workers, which means that $\delta = 1+r_p$ in equation (3). Hence if a bank charges interest r , the actual rate of its loans is $R = (1+r)/(1+r_p)$. At the optimum, all banks offer $R = \hat{R}$, the market clearing rate. It follows that

$$1+r = \hat{R}(1+r_p).$$

That is, banks mark up the interest rate on loans to pass the cost of deposit due to the interest rate policy on to entrepreneurs. If a bank issues notes of aggregate face value D , then at $t = 1$, with the interest rate policy, its liability to the note holders is $D(1+r_p)$. The equity value in the bad state is thus $G + \underline{Y} - D(1+r_p)$, which, by the borrowing constraint, cannot be smaller than $(1-\mu)G$. With $\underline{Y} = \frac{A(1+r)}{A\alpha}D$, this constraint is equivalent to

$$D(1+r_p) \leq \mu G + \frac{A(1+r)}{A\alpha}D, \tag{30}$$

which, with $1 + r = \widehat{R}(1 + r_p)$, is equivalent to

$$D \leq \frac{\mu G}{1 + r_p} \left(1 - \frac{A}{A\alpha} \widehat{R}\right)^{-1} := \overline{D}^* (\widehat{R}; G).$$

According to proposition 1, if $\widehat{R} = R^{SB}$, the profit margin of lending is zero and banks are indifferent in any D , while if $\widehat{R} > R^{SB}$, the profit margin is positive and banks want to lend as much as they can, that is, the above borrowing constraint is binding. Hence, the supply of notes is

$$D^* (\widehat{R}; G) = \left\{ \begin{array}{l} 0 \text{ if } \widehat{R} < R^{SB} \\ [0, \overline{D} (\widehat{R}; G)] \text{ if } \widehat{R} = R^{SB} \\ \overline{D} (\widehat{R}; G) \text{ if } \widehat{R} > R^{SB} \end{array} \right\}.$$

Observe that in equation $E^* (\widehat{R}) = \overline{D}^* (\widehat{R}; G)$, the factor $1 + r_p$ is canceled out, and the equation is equivalent to $E (\widehat{R}) = \overline{D} (\widehat{R}; G)$, where $E (\cdot)$ and $\overline{D} (\cdot)$ are the demand and maximum supply functions of notes in proposition 3, namely in the absence of the policy. It follows that $E^* (\widehat{R}) = D^* (\widehat{R}; G)$ if and only if $E (\widehat{R}) = D (\widehat{R}; G)$, where $D (\cdot)$ is the supply functions of notes in proposition 3. That is, the equilibrium \widehat{R} in the presence of the interest rate policy is the same as that in its absence. Therefore, the interest rate policy produces no real effects.

- (ii) In the circumstance where the nominal wage of workers stays at w invariant to the policy rate r_p , if an entrepreneur borrows notes of face value E , he still hires E/w (rather than $E(1 + r_p)/w$) workers. This implies (a) that the aggregate demand for notes is wL (rather than $wL/(1 + r_p)$), which is equalized to the aggregate supply D in equilibrium; and (b) that the discount factor in (3) is $\delta = 1$, whereby the actual interest rate to entrepreneurs is $R = 1 + r$ (rather than $R = (1 + r)/(1 + r_p)$). Given \widehat{R} , $L = (\frac{\overline{A}\alpha}{w})^{\frac{1}{1-\alpha}} \widehat{R}^{\frac{-1}{1-\alpha}}$ by (5). It follows from implication (a) that $D = w(\frac{\overline{A}\alpha}{w})^{\frac{1}{1-\alpha}} \widehat{R}^{\frac{-1}{1-\alpha}}$.

At the optimum all banks choose $R = \widehat{R}$, which together with implication (b) means $1 + r = \widehat{R}$. Substitute these into (30) and banks' borrowing constraint becomes

$$w \left(\frac{\overline{A}\alpha}{w\widehat{R}} \right)^{\frac{1}{1-\alpha}} \left[(1 + r_p) - \frac{A}{A\alpha} \widehat{R} \right] \leq \mu G. \tag{31}$$

This borrowing constraint is binding if the profit margin of lending is positive, as before. Given that the workers' real wage is now $w(1 + r_p)$, the profit to a bank from lending to one entrepreneur becomes $A_e L^\alpha - w(1 + r_p)L - V$. With L and V as functions of $R = \widehat{R}$ given by (5) and (6), the profit margin with sticky wage is $[\overline{A}\alpha/(w^\alpha \widehat{R})]^{\frac{1}{1-\alpha}} [\widehat{R}/R^{SB} - (1 + r_p)]$. This profit margin never goes below 0. Therefore,

$$1 + r_p \leq \widehat{R}/R^{SB}, \tag{32}$$

which implies that $\widehat{R} \geq R^{SB}$.

The equilibrium actual interest rate \widehat{R} is pinned down by conditions (31) and (32), and the fact that one of them must be binding: if (32) is not binding—that is, if the profit margin of issuance is positive—then banks keep issuing notes until the no-default constraint, (31), is binding. Note that (31) is equivalent to

$$1 + r_p \leq \Phi(\widehat{R}), \tag{33}$$

where

$$\Phi(R) \equiv \frac{\mu G w^{\frac{\alpha}{1-\alpha}}}{(\overline{A}\alpha)^{\frac{1}{1-\alpha}}} R^{\frac{1}{1-\alpha}} + \frac{A}{A\alpha} R, \tag{34}$$

and obviously $\Phi'(\cdot) > 0$, and the inverse function, $\Phi^{-1}(\cdot)$, exists. Then in equilibrium, inequalities (33) and (32) hold, and one of them must be binding. Therefore,

$$1 + r_p = \min \left(\widehat{R}/R^{SB}, \Phi(\widehat{R}) \right) := H(\widehat{R}). \tag{35}$$

We only need to consider function $H(\cdot)$ over $\widehat{R} \geq R^{SB}$.

CLAIM P6. Over $\widehat{R} \geq R^{SB}$, if $G \geq \frac{1}{\mu}G^*$, $H(\widehat{R}) = \widehat{R}/R^{SB}$; and if $G < \frac{1}{\mu}G^*$, then

$$H(\widehat{R}) = \left\{ \begin{array}{l} \Phi(\widehat{R}) \text{ if } \widehat{R} \leq R^* \\ \widehat{R}/R^{SB} \text{ if } \widehat{R} \geq R^* \end{array} \right\},$$

where

$$R^* = \frac{\overline{A}\alpha}{w} \left[\frac{q(\overline{A}\alpha - \underline{A})}{\mu G} \right]^{\frac{1-\alpha}{\alpha}}. \tag{36}$$

Proof. Let $\chi(R) \equiv \Phi(R) - R/R^{SB}$ for $R \geq 0$. Then $\chi(R) = 0$ has two roots: 0 and R^* . Then (a) $R^* \geq R^{SB}$ if and only if $G \leq \frac{1}{\mu}G^*$; (b) $\chi < 0$ for $R \in (0, R^*)$ and $\chi > 0$ for $R > R^*$, because $\chi'(0) < 0$. These two results lead to the claim. Q.E.D.

With help of the claim, we solve \widehat{R} as a function of r_p from (35) as follows. If $G \geq \frac{1}{\mu}G^*$, $\widehat{R} = R^{SB}(1 + r_p)$. If $G < \frac{1}{\mu}G^*$, then

$$\widehat{R} = \left\{ \begin{array}{l} \Phi^{-1}(1 + r_p) \text{ if } r_p \leq r_p^* \\ R^{SB}(1 + r_p) \text{ if } r_p \geq r_p^* \end{array} \right\},$$

where $r_p^* \equiv R^*/R^{SB} - 1$, which, as $\chi(R^*) = 0$, also equals $\Phi(R^*) - 1$; hence $r_p^* \geq 0 \Leftrightarrow R^* \geq R^{SB} \Leftrightarrow G \leq \frac{1}{\mu}G^*$.

In both cases, \widehat{R} increases with the policy rate r_p to infinity. And at $r_p = 0$, namely in the absence of the CB's intervention, $\widehat{R} < R^{FB}$ in this circumstance of banks over-lending. Therefore, there exists a unique policy rate at which $\widehat{R} = R^{FB}$ and the first-best allocation is attained.

Banks obtain zero profit if the non-negative profit constraint, (32), is binding, namely $H(\widehat{R}) = \widehat{R}/R^{SB}$, which, by claim P6, is the case if $\widehat{R} \geq R^*$. With the optimal policy rate, $\widehat{R} = R^{FB}$, and hence banks obtain zero profit if $R^{FB} \geq R^*$, which is equivalent to $G \geq G_s := \frac{1}{\mu}[(q\overline{A}\alpha + (1-q)\underline{A}) - \underline{A}](A_e\alpha/w)^{\frac{\alpha}{1-\alpha}}$. It is straightforward to check that $G^{FB} < G_s < \frac{1}{\mu}G^*$. Q.E.D.

Proof of Proposition 7

According to the discussion of subsection 3.2, $\overline{Y} = D(1 + r)$ and $\underline{Y} = \underline{A}/(\overline{A}\alpha) \times D(1+r)$. Therefore, $Y = [q+(1-q)\underline{A}/(\overline{A}\alpha)] \times D(1+r)$.

Moreover, at the optimum, banks all offer $1 + r = \widehat{R}$. It follows that if banks issue D , the capital adequacy ratio is

$$c = \frac{G + \{[q + (1 - q)\underline{A}/(\overline{A}\alpha)]\widehat{R} - 1\}D}{G + [q + (1 - q)\underline{A}/(\overline{A}\alpha)]\widehat{R}D}.$$

If banks' issuance is at the first quantity and the first-best allocation is attained, then $D = D^{FB}$ and $R = R^{FB}$ in the above formula and the capital adequacy ratio equals

$$c^{FB} := \frac{A_e\alpha G + (1 - q)\underline{A}(1 - \alpha)D^{FB}}{A_e\alpha G + [q\overline{A}\alpha + (1 - q)\underline{A}]D^{FB}}.$$

The regulation restricts $c \geq c^{FB}$. To prove the proposition, it suffices to show that this restriction is equivalent to $D \leq D^{FB}$, the constraint which, as we saw in the main text, will lead to the first-best allocation. This equivalence follows from two observations: (i) $c = c^{FB}$ if $D = D^{FB}$ and (ii) c decreases with D because $c = 1 - \frac{D}{G + [q + (1 - q)\underline{A}/(\overline{A}\alpha)]\widehat{R}D} := c(D, \widehat{R})$ and moreover, $\frac{\partial c}{\partial D} < 0$ and $\frac{\partial c}{\partial \widehat{R}} > 0$ while \widehat{R} decreases with D .

Q.E.D.

References

- Allen, F., and D. Gale. 1998. "Optimal Financial Crises." *Journal of Finance* 53 (4): 1245–84.
- Brunnermeier, M., and Y. Sannikov. 2013. "The I-Theory of Money." Working Paper, Princeton University.
- Cameron, R., O. Crisp, H. T. Patrick, and R. Tilly. 1967. *Banking in the Early Stages of Industrialization*. Oxford University Press.
- Cavalcanti, R., A. Erosa, and T. Temzelides. 1999. "Private Money and Reserve Management in a Random-Matching Model." *Journal of Political Economy* 107 (5): 929–45.
- Chen, S.-H. 2018. "The Credit Channel Transmission Mechanism and the Nonlinear Growth and Welfare Effects of Inflation and Taxes." *Economic Inquiry* 56 (2): 724–44.
- Dang, T., G. Gorton, and B. Holmstrom. 2012. "Ignorance, Debt and Financial Crises." Mimeo.

- Diamond, D., and R. Rajan. 2006. "Money in a Theory of Banking." *American Economic Review* 96 (1): 30–53.
- Donaldson, J. D., G. Piacentino, and A. Thakor. 2018. "Warehouse Banking." *Journal of Financial Economics* 129 (2): 250–67.
- Gertler, M., and P. Karadi. 2011. "A Model of Unconventional Monetary Policy." *Journal of Monetary Economics* 58 (1): 17–34.
- Gertler, M., and N. Kiyotaki. 2010. "Financial Intermediation and Credit Policy in Business Cycle Analysis." In *Handbook of Monetary Economics*, Vol. 3A, ed. B. M. Friedman and M. Woodford, 547–99 (chapter 11). North-Holland.
- Goodfriend, M., and B. McCallum. 2007. "Banking and Interest Rates in Monetary Policy Analysis: A Quantitative Exploration." *Journal of Monetary Economics* 54 (5): 1480–1507.
- Gorton, G. 1999. "Pricing Free Bank Notes." *Journal of Monetary Economics* 44 (1): 33–64.
- Gorton, G., and G. Pennacchi. 1990. "Financial Intermediaries and Liquidity Creation." *Journal of Finance* 45 (1): 49–71.
- Gorton, G., and A. Winton. 2003. "Financial Intermediation." In *The Handbook of the Economics of Finance: Corporate Finance*, Vol. 1, ed. G. Constantinides, M. Harris, and R. Stulz, 431–552 (chapter 8). Amsterdam: Elsevier Science.
- Hart, O., and L. Zingales. 2015. "Liquidity and Inefficient Investment." *Journal of the European Economic Association* 13 (5): 737–69.
- Holmstrom, B., and J. Tirole. 1997. "Financial Intermediation, Loanable Funds, and the Real Sector." *Quarterly Journal of Economics* 112 (3): 663–91.
- Jakab, Z., and M. Kumhof. 2015. "Banks Are Not Intermediaries of Loanable Funds — And Why This Matters." Working Paper No. 529, Bank of England.
- Jensen, M. C., and W. H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–60.
- Kiyotaki, N., and J. Moore. 2001. "Evil is the Root of All Money." Clarendon Lecture No. 1, Oxford University.
- Kreps, D., and J. Scheinkman. 1983. "Quantity Precommitment and Bertrand Competition Yields Cournot Outcomes." *Bell Journal of Economics* 14 (2): 326–37.

- Mishkin, F. S. 2016. *The Economics of Money, Banking, and Financial Markets*. 11th edition. Pearson Education.
- Monnet, C., and D. Sanches. 2015. "Private Money and Banking Regulation." *Journal of Money, Credit and Banking* 47 (6): 1031–62.
- Samuelson, P. 1958. "An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money." *Journal of Political Economy* 66: 467–82.
- Stein, J. 2012. "Monetary Policy as Financial Stability Regulation." *Quarterly Journal of Economics* 127 (1): 57–95.
- Tobin, J. 1963. "Commercial Banks as Creators of Money?" In *Banking and Monetary Studies*, ed. D. Carson, for the Comptroller of the Currency, U.S. Treasury. Richard D. Irwin.
- Wang, T. 2010. "Risk Sharing, (Over)Leverage, and the Dimensions of Competition." Mimeo.
- . 2015. "Competition and Increasing Returns to Scale: A Model of Bank Size." *Economic Journal* 125 (585): 989–1014.
- Williamson, S. 1999. "Private Money." *Journal of Money, Credit and Banking* 31 (August): 469–91.