Universidad de Valladolid

Facultad de Ciencias Económicas y Empresariales

Trabajo Fin de Grado

Grado en Economía

Beyond the Central Bank: A Formal Approach to Free Banking Theory

Presentado por:

Alejo Hubble Villoslada

Tutelado por:

Luis Pablo de la Horra Ruiz

Valladolid, 18 de junio de 2025

"It is not from the benevolence of the butcher, the brewer, or the baker that we expect our dinner,
but from their regard to their own interest."

— **Adam Smith**, *The Wealth of Nations* (1776)

Beyond the Central Bank: A Formal Approach to Free Banking Theory

Alejo Hubble

June 19, 2025

Abstract

This paper develops a parsimonious model of bank behavior under a free banking regime, where the issuance of inside money is not centrally managed but instead governed by market forces. Building on a partial equilibrium framework influenced by Selgins work on Free Banking Theory, I show that a profit-maximizing bank will endogenously limit the issuance of demandable liabilities due to rising liquidity costs associated with reserve depletion. This result provides a microeconomic foundation for the theory of monetary equilibrium, whereby the supply of inside money adjusts to match demand ($S_m = D_m$). At the macro level, I explore the implications of this equilibrium condition for the price level and the natural interest rate, arguing that a free banking system can achieve monetary stability without central intervention. While the model abstracts from frictions such as coordination failures or asymmetric information, it offers a tractable foundation for understanding self-regulation in competitive banking environments. The findings challenge conventional views that associate free banking with monetary instability, and suggest that under certain conditions, decentralized banking can deliver disciplined and efficient outcomes.

Keywords: Free Banking, Liquidity Constraints, Profit Maximization, Monetary Equilibrium.

JEL Codes: E42, E52 ,G21, E58.

Contents

List of Figures

1 Introduction

Banks occupy a foundational role in modern economies by enabling the transformation of savings into productive investment while simultaneously providing liquidity to depositors. At the heart of this dual function lies the process of maturity transformation: banks hold long-term, illiquid assets while issuing short-term, liquid liabilities. This tradeoff between illiquidity and liquidity provision is elegantly formalized in the influential model by Diamond and Dybvig (1983). In their framework, agents face uncertainty about their future liquidity needs. In the absence of financial intermediaries, this uncertainty leads to inefficient investment in short-term, low-return assets. The introduction of banks, however, allows for efficient risk-sharing: individuals can deposit funds in institutions that invest in long-term projects while preserving access to liquidity in the event of early withdrawal. Thus, banks emerge as a welfare-enhancing solution to a fundamental problem of intertemporal allocation under uncertainty.

Historically, the emergence and regulation of banking have been accompanied by deep theoretical and political controversies. Chief among these was the 19^{th}-century debate between the Currency School and the Banking School. The Currency School advocated for a strict link between the issuance of banknotes and gold reserves, fearing that unconstrained note issue would generate inflation and financial instability. In contrast, the Banking School emphasized the role of endogenous demand for money and posited that overissue would be self-correcting due to the “law of reflux,” whereby excess notes would be returned to issuing banks for redemption. This theory suggested that competitive forces and the clearing mechanism between banks naturally regulate the quantity of money in circulation.

In retrospect, the Currency School largely won the intellectual and institutional battle. Central banks were established to centralize note issuance, manage interest rates, and stabilize prices. Yet despite their predominance, central banks have exhibited several shortcomings: First, central banks can create moral hazard by acting as lenders of last resort, encouraging excessive risk-taking (Dowd 1996; Anginer and Demirgüç-Kunt 2018)[1]. Second, their political exposure may lead to inflation bias or policy driven by short-

[1] The identity, under certain circumstances, of deposit insurance and discount window is consistent with Diamond and Dybvig (1983).

term interests[2]. Third, from Austrian and free banking perspectives, central banks are seen as distorting natural credit markets and suppressing monetary competition. Others argue that central banks suffer from knowledge limitations, making it impossible to set interest rates optimally or to distinguish insolvency from liquidity needs. Additional critiques include their role in widening inequality through asset purchases, creating market dependence on low rates and liquidity, and functioning as central planners that disrupt price signals.

Given that the financial system operates as the heart and veins of the entire economy—circulating credit, coordinating investment, and maintaining liquidity—these potential shortcomings deserve serious consideration. If the institutional framework responsible for such vital functions is flawed or suboptimal, the consequences extend far beyond the banking sector itself, potentially affecting growth, stability, and the allocation of resources throughout the economy. Engaging critically with these issues is thus not only an academic exercise, but a necessary step toward understanding whether alternative arrangements could offer more resilient and efficient outcomes. The potential limitations outlined and the importance of the topic invite a reconsideration of alternatives—among them, free banking.

Free banking refers to a regime in which the issuance of inside money (such as deposits and banknotes) is decentralized and governed by competitive forces, rather than a monopolistic central authority. While proponents of free banking—such as Selgin (1988b) and White—have articulated its theoretical merits and supported them with historical examples, formal attempts to model bank behavior under such a system remain scarce.

The purpose of this paper is not to adjudicate the broader normative debate over the desirability or correctness of free banking theory. Rather, it aims to contribute to the theory by formally modeling the behavior of a profit-maximizing bank in a competitive, unregulated environment. Building on Selgin's conceptual foundation, chiefly Selgin (1988b) and Selgin (1988a), this

[2] Alesina (1988) undertake an empirical and theoretical examination of how political pressures affect macroeconomic outcomes. They find a statistically significant inverse relationship between central bank independence and average inflation rates across countries. Furthermore, nations with more independent central banks tend to experience lower average fiscal deficits and less fiscal volatility. These findings support the idea that political influence over monetary policy can lead to inflationary bias, highlighting the risk of politicized central banking.

paper develops a microeconomic model in which banks face liquidity costs that increase nonlinearly as reserves are depleted. Within this framework, I demonstrate that banks have endogenous incentives to limit the issuance of liabilities, aligning money supply with money demand.

While the model is not constructed to advocate for or against free banking as a policy regime, its results lend support to the internal logic of the theory by showing that monetary equilibrium can arise from decentralized profit-maximizing behavior. In this sense, the model strengthens the plausibility of free banking by embedding its core mechanisms within the familiar apparatus of neoclassical optimization. Moreover, it offers a foundation for future empirical investigations that could assess the historical validity of free banking regimes or simulate their outcomes under varying institutional constraints.

This paper is structured as follows. Section 2 presents a review of the relevant literature. Section 3 introduces the theoretical model developed for this study, outlining its key assumptions, internal mechanics, and core insights. In Section 4, I analyze the macroeconomic implications that emerge from the model's monetary equilibrium, particularly in terms of stability, liquidity provision, and policy relevance. Finally, Section 5 concludes by summarizing the findings and discussing their broader theoretical and policy implications.

2 Literature Review

Despite the rich historical and theoretical discussion surrounding free banking, there have been remarkably few attempts to formally model how a bank would operate under such a system. Much of the literature on free banking—such as that by Selgin, White, and Dowd—has focused on conceptual or institutional analysis rather than on rigorous microeconomic foundations. One of the rare exceptions is the work by Cavalcanti, Erosa, and Temzelides Cavalcanti, Erosa, and Temzelides (1999), who develop a random-matching model in which banks issue private liabilities that circulate as media of exchange. Their model shows that note redemption acts as a force that is sufficient to stabilize note issue by the banking sector. While this is an important insight, it does not explicitly demonstrate how bank note issue restriction can be in line with bank profit-maximization.

A more recent contribution is provided by Sanches (2016), who develops a general equilibrium model in which the banking system may either function smoothly or collapse. These outcomes arise endogenously from agent's expectations and information, without requiring coordination failures. The collapse scenario hinges on two assumptions that warrant scrutiny.

First, the model assumes depositors lack information about the quality of bank collateral. In a competitive free banking regime, however, such opacity is less plausible. Banks would likely face strong reputational pressures to disclose asset quality, with transparency enforced through market signals such as equity prices, third-party audits, and risk-based pricing by insurers.

Second, the model assumes that expectations of deteriorating monetary conditions persist over time, effectively creating a liquidity trap in which even higher deposit returns—including through instruments like lotteries for notes returns—fail to attract demand. While theoretically interesting, the empirical relevance of such traps is debatable (Selgin 1988b).

Another insightful contribution is offered by Azariadis and Kaas (2023), who present a dynamic general equilibrium model in which privately issued liabilities circulate in the absence of government fiat money. Their framework elegantly demonstrates how heterogeneity in agent reputations leads to equilibrium configurations where marginal rates of substitution (MRS) differ across agents, reflecting a constrained non-Pareto optimal outcome. Moreover, their model explains how persistent premiums and discounts emerge endogenously, generating volatility in the use of private monies.

Several of the assumptions behind the model may be open to debate. In particular, the lack of private insurance mechanisms for lenders, and the omission of interbank incentives to accept liabilities at par—mechanisms emphasized by free banking theorists such as Selgin—limit the applicability of the results. In models with institutional arrangements that encourage par acceptance (e.g., through reserve arbitrage), such price dispersions may be arbitraged away, making the system function more like one based on a common outside money. Thus, while Azariadis and Kaas (2023) provide a compelling benchmark, it leaves open the question of whether such volatility is an unavoidable feature of private money systems or a product of restrictive assumptions.

Another important foundation for understanding banking fragility is provided by the seminal model of Diamond and Dybvig (1983). Their framework illustrates how banks that engage in maturity transformation—converting short-term deposits into long-term loans—can be vulnerable to self-fulfilling runs. While not situated in a free banking context, their analysis is central to understanding the inherent risks banks face when providing liquidity to consumers with uncertain consumption timing. The possibility of runs justifies the existence of deposit insurance and lender-of-last-resort interventions, which stand in contrast to the self-regulatory mechanisms emphasized in free banking theory.

A recent contribution that connects more directly to macroeconomic policy is offered by Salter and Young (2018), who construct a model of free banking to examine whether such a system could stabilize nominal GDP (NGDP) growth. Their analysis provides a bridge between the microeconomic structure of free banking and its potential to support macroeconomic stability goals, such as maintaining a stable price level. This connection is particularly relevant for the discussion later in this paper, where we assess the price level consequences of a free banking system.

The model developed in this paper adds a complementary perspective to all this literature by showing that, under competitive conditions, profit-maximizing banks will endogenously limit the issuance of liabilities in response to demand. This provides a formal microfoundation for self-regulation in a free banking regime and offers a clearer understanding of how market discipline operates in such systems.

3 Partial equilibrium model

In this section, I attempt to formalize the behavior of a profit maximizing bank in partial equilibrium under a free banking system.

3.1 Partial equilibrium behavior under fixed money demand

Firstly, I assume that the bank's total demandable liabilities Q_l—composed of demand deposits and circulating notes—are equal to the demand for inside money D_m[3]. It is important to clarify this key modeling assumption: the

[3] "The demand for money, properly understood, refers to the desire to hold money as part of a financial portfolio. A bank borrower contributes no more to the demand for

model presumes that any issuance of liabilities (Q_l) by the bank that exceeds the internal demand (D_m) does not remain within the bank but instead flows immediately to other institutions in the system. In other words, once $Q_l > D_m$, the excess liabilities are not retained by the bank but are either spent or redeemed elsewhere. This justifies the working assumption that $Q_l = D_m$ for the individual bank. Such behavior is consistent with established bank theory, which holds that liabilities not desired by the public or other banks cannot be passively held by the issuing bank itself. These dynamics effectively constrain liability creation to the quantity that can be absorbed by the market. A detailed treatment of the consequences for reserves and interbank flows will be developed below.

D_m is also treated as externally given to individual banks, which is a standard feature of models of banking under competition. I will also assume that D_m is constant (I'll be relaxing this assumption further on).

$$Q_l = D_m = \text{constant}$$

The bank chooses the composition of its balance sheet to maximize profit, subject to the constraint that total liabilities must equal the total value of interest-earning assets Q_a–which we interpret as loans extended to the public– and reserves R. $Q_a + R = Q_l$. In this model, I define reserves as any form of base money that is widely accepted by banks for the purpose of interbank clearing (it is assumed that the amount of R in the aggregate bank system is constant). While in modern economies this typically refers to central bank-issued fiat money (e.g., reserve balances or physical currency), the concept is more general. Under a free banking regime, reserves could consist of commodity money such as gold or silver. In a hypothetical scenario without a central bank, widely accepted banknotes from a previous CB could potentially serve as reserves. Similarly, in the future, a highly liquid and universally trusted cryptocurrency—such as Bitcoin—could, in principle, fulfill the same role. Another posibilitie could be the liabilities of a private bank that performs the functions of a central bank (even if it does not hold the monopoly on issuance). The essential feature of reserves in this model is

money than a ticket agent contributes to the demand for plays and concerts; only holders of money or actual occupants of concert seats contribute to demand" (Selgin (1988b), p. 61).

not their origin but their acceptability and reliability for settling interbank claims.

The balance sheet identity $Q_a + R = Q_l$ follows naturally from treating the bank as a pure financial intermediary: all issued liabilities are backed either by loans or by reserves held to meet redemption demands or adverse clearing[4].

Bank profit is given by:

$$\Pi = f(\theta) + \text{IPA} - \text{IPL} - C_L$$

where:

- $f(\theta)$ **(Other Sources of Net Income)**: represents a reduced-form function that captures the net income from sources not explicitly modeled in the intermediation process. These include fee-based revenues, equity capital returns, and other operational surpluses or expenditures. In practice, banks earn profits from non-intermediated activities such as service charges, investment income, or payment services. Moreover, equity capital contributes to funding without requiring interest payments. Including $f(\theta)$ ensures that the model reflects a more complete and realistic depiction of bank profitability beyond traditional deposit-loan intermediation.

- **IPA (Interest-earning Present Assets)**: represents the present value of the bank's interest-earning assets. It is defined as:

  $$\text{IPA} = Q_a \cdot i_a$$

  with i_a being the interest rate earned on the bank's interest-earning assets.

[4]Selgin's account of adverse clearings in a free banking system mirrors the traditional understanding of excess reserve discipline in orthodox banking theory: banks that overissue liabilities face a loss of reserves through interbank clearings. However, unlike central bank-regulated systems, this disciplining effect arises endogenously in a free banking regime due to the public's and rival banks' preference for more reputable and reliably redeemable notes. This assumes a level of note discrimination by the public that enforces competitive restraint. See Selgin (1988b), esp. Ch. 6–7.

- **IPL (Interest-bearing Present Liabilities)**: corresponds to the present value of the bank's liabilities (just demand deposits in this case) on which it pays interest:

$$\text{IPL} = Q_l \cdot i_l$$

  where Q_l is the total quantity of liabilities (inside money issued), and i_l is the interest paid on them.

- **Liquidity Costs** C_L: capture the expected cost of meeting clearing obligations with reduced reserves. These costs are modeled as a function of the inverse reserve ratio. The liquidity cost function employed in this model takes the form:

$$C_L = \alpha \left(\frac{Q_l}{R} \right)^{\gamma},$$

  where $\alpha > 1$ and $\gamma > 2$. The rationale for this functional form is twofold. First, the exponent γ captures the idea that liquidity costs rise more than proportionally as the ratio of liabilities to reserves increases. That is, as a bank stretches its reserves thinner to support greater levels of liabilities, the risk and associated costs of maintaining liquidity grow at an accelerating rate. This convexity is essential to reflect the real-world dynamics of liquidity stress and the potential for non-linear responses in clearing systems, interbank markets, and depositor behavior. For instance, depositors may tolerate relatively high liability-to-reserve (Q_l/R) ratios up to a certain threshold, perceiving the bank as sufficiently liquid. However, once this threshold is crossed, even modest increases in the ratio can provoke disproportionate reductions in deposit demand, as confidence erodes. This adjustment need not take the form of a run; it may simply manifest as a shift in preferences toward more liquid institutions.

  Second, the scaling parameter α ensures that the cost function has the correct monetary dimensions, allowing it to be meaningfully integrated into the profit function. It also represents institutional or market-specific features such as the frictions in interbank clearing, premiums on deposit insurances, and the availability or cost of emergency liquidity support. While both α and γ may vary depending on the structure and robustness of the banking system under consideration, the qualitative conclusions derived from the model remain robust as long as $\alpha > 1$ and $\gamma > 2$, ensuring that liquidity costs are strictly increasing and economically relevant over the relevant domain.

$$C_L = \alpha \left(\frac{Q_l}{R}\right)^{\gamma}, \quad \text{with} \quad f' > 0$$
$$\alpha > 1, \quad \gamma > 2$$

From the above, it follows that:

$$\Pi = f(\theta) + Q_a \cdot i_a - Q_l \cdot i_l - \alpha \left(\frac{Q_l}{R}\right)^{\gamma}$$

Using the balance sheet constraint $Q_l = Q_a + R$, we can express reserves as:

$$R = Q_l - Q_a = D_m - Q_a$$

$$R > 0$$

therefore:

$$Q_l - Q_a > 0$$

Substituting this into the profit function, we can write profit as a function of Q_a alone:

$$\Pi = f(\theta) + Q_a \cdot i_a - D_m \cdot i_l - \alpha \left(\frac{D_m}{D_m - Q_a}\right)^{\gamma}$$

In this profit function, the only endogenous variable from the perspective of the individual bank is Q_a. The interest rates i_a and i_l are assumed to be set by the market and thus taken as given by the bank. Similarly, D_m is treated as an exogenous parameter, reflecting the prevailing demand for inside money. Consequently, in order to understand the bank's profit-maximizing behavior, we derive the first-order condition of profit with respect to Q_a.

To determine the bank's optimal level of interest-earning assets Q_a, we differentiate the profit function with respect to Q_a and set it equal to zero ($f(\theta)$ is assumed for simplicity not to depend on Q_a) :

$$\frac{d\Pi}{dQ_a} = i_a - \alpha \cdot \gamma \left(\frac{D_m}{D_m - Q_a} \right)^{\gamma - 1} \cdot \frac{D_m}{(D_m - Q_a)^2}$$

The first-order condition is:

$$i_a - \alpha \cdot \gamma \left(\frac{D_m}{D_m - Q_a} \right)^{\gamma - 1} \cdot \frac{D_m}{(D_m - Q_a)^2} = 0$$

Solving for i_a gives

$$i_a = \alpha \cdot \gamma \left(\frac{D_m}{D_m - Q_a} \right)^{\gamma - 1} \cdot \frac{D_m}{(D_m - Q_a)^2}$$

This condition states that the bank expands its portfolio of interest-earning assets Q_a up to the point where the marginal revenue from lending (i_a), equals the marginal cost of reduced liquidity. As Q_a increases, reserves $R = D_m - Q_a$ fall, raising liquidity costs and eventually offsetting the gains from additional interest income. In contrast, modern banking systems—particularly since the 2008 financial crisis—operate under a floor reserve regime, wherein reserves are no longer scarce, and monetary policy is implemented primarily through administered interest rates. In such systems, the central bank supplies ample reserves, and short-term interest rates are controlled via the interest paid on reserve balances (IORB) in the United States or the marginal deposit facility rate in the euro area. As a result, reserve holdings no longer constitute a binding constraint on bank lending. Instead, contemporary lending behavior is shaped more directly by considerations such as credit risk, capital adequacy, and yield optimization, rather than by liquidity management concerns rooted in reserve depletion.

The structure of the partial equilibrium model is consistent with standard profit maximization theory. The first-order condition derived from the profit function ensures that the marginal revenue from increasing interest-earning assets Q_a equals the marginal increase in liquidity costs. This condition characterizes an equilibrium point: if the bank were to increase Q_a beyond this point, the associated rise in liquidity costs—due to a reduction in reserves—would outweigh the additional interest income, leading to a decline in profits. Thus, the model captures a natural upper bound on asset expansion and reinforces the theoretical soundness of the equilibrium configuration.

Notably, the model's liquidity cost function exhibits a steep rise as Q_a increases, the marginal cost of extending additional loans becomes prohibitively high. This sharp increase can be interpreted as capturing the latent pressure of a potential bank run. Thus, the model indirectly reflects the risk environment that would trigger runs, and suggests that banks are disciplined to avoid such a zone by maintaining prudent reserve ratios. This reinforces the self-regulating logic central to the theory of free banking.

3.2 Partial equilibrium behavior with changing money demand

We now extend the partial equilibrium setting by allowing for changes in the demand for inside money D_m faced by an individual bank. While the initial setup held D_m constant, we now treat it as a variable exogenous to the bank's portfolio decision but relevant for its liquidity constraint.

From the first-order condition previously derived, we can observe that an increase in D_m reduces the marginal liquidity cost, as the derivative of the liquidity cost function $\frac{\partial C_L}{\partial Q_a}$ declines with a higher D_m. This relaxation of liquidity pressure incentivizes the bank to expand its interest-earning assets (Q_a). Conversely, a fall in D_m increases the marginal liquidity cost and leads the bank to contract (Q_a). The model thus provides a coherent behavioral response to shifts in money demand and reinforces the idea that even in its parsimonious form, the framework captures core features of profit-maximizing behavior under liquidity constraints in a free banking environment.

3.3 Partial equilibrium model insights

All the above reveals that, under equilibrium, the quantity of interest-earning assets (Q_a) is constrained by the marginal liquidity costs arising from reserve depletion. This captures a fundamental insight regarding the regulation of inside-money supply under free banking. As Selgin (1988b) explains, it is not just the costs connected with the issue of inside money which regulate its supply under free banking; rather, it is these costs plus the costs associated with the return of notes and checks to their issuers for redemption in base money, that is, liquidity costs.

When there is an exogenous increase in the demand for inside money (D_m), the resulting decline in liquidity costs incentivizes banks to expand (Q_a), which in turn raises their liabilities (Q_l)[5] and maintains the accounting identity $Q_l = Q_a + R$.

However, in the absence of a change in demand, an increase in Q_a must be financed by a reduction in reserves or Q_a, since liabilities cannot sustainably exceed demand. After a brief expansion in Q_l, interbank clearing mechanisms force the bank to contract reserves or reduce Q_a back to the previous level (likely incurring some losses due to the early termination of assets), thus restoring equilibrium.

For further insight, we solve the first-order condition for Q_a and differentiate the resulting expression with respect to D_m[6]:

$$Q_a = D_m - \left(\frac{\alpha \cdot \gamma \cdot D_m^{\gamma}}{i_a}\right)^{\frac{1}{\gamma+1}}$$

Thus, the sensitivity of the optimal asset choice to changes in money demand is given by

$$\frac{dQ_a}{dD_m} > 0,$$

indicating that an increase in demand for inside money(D_m) leads the bank to expand its portfolio of interest-earning assets.[7].

Figure 1 illustrates the core qualitative insight of the model developed in this section. The figure represents the profit function vs Q_a. The term $f(\theta)$ is omitted in this graph, as it is unknown and does not affect the qualitative shape of the function. For the sake of plotting, the parameters are assumed as follows: $\alpha = 50$, $\gamma = 39$, $D_m = 50$ million, $i_l = 0.01$, and $i_a = 0.0475$[8]. The plot captures how, under a free banking system, a profit-maximizing bank

[5] This simultaneous increase in assets and liabilities—such as when a commercial bank extends a loan and thereby creates a corresponding deposit—is an accounting necessity. The process of loan creation inherently expands both sides of a bank's balance sheet, with the loan recorded as an asset and the deposit as a liability. In this case an increase in (Q_a) and (Q_l) are both sides of the same coin (McLeay, Radia, and Thomas 2014).

[6] see Appendix A.

[7] Except for extreme values of D_m and Q_a, where the absolute difference of D_m respect to D_m-Q_a makes the liquidity cost function approach infinite.

[8] The interest spread $i_a - i_l = 3.75\%$ represents the average US NIM from 1986 to 2020 (Federal Financial Institutions Examination Council (US) and Federal Reserve Bank of St. Louis 2020).

chooses a finite and internal optimum for the amount of loans extended, given the demand for inside money. This outcome underscores the central theoretical contribution of the present discussion: banks do not expand credit without bound, but instead face an optimal lending threshold constrained by liquidity costs.

This insight is particularly relevant in the light a traditional criticism of free banking: that in the absence of central coordination, banks would overissue money (whether in the form of banknotes or deposits), generating inflation and macroeconomic imbalances. For instance, John Maynard Keynes argued that without a central authority to oversee and moderate credit creation, banks might be incentivized to expand their liabilities excessively during periods of economic optimism, thereby amplifying aggregate demand beyond sustainable levels and triggering inflationary pressures.

Although the shape and underlying logic of the profit function are convincing, two quantitative limitations seem evident in the figure, namely negative profits and a ratio of Q_a to D_m of just 16.27%, implying that reserves make up 83.73% of liabilities—an unrealistically high reserve ratio for most real-world banking systems[9]. Importantly, these do not compromise the qualitative validity of the model. First, the profit values are negative under the chosen parameter settings. This outcome is a direct consequence of the specific parameter selection. Reducing the values of α and γ slows the growth of liquidity costs, thereby permitting positive profits. Additionally, lower α and γ values allow for a larger Q_a relative to D_m before liquidity costs heavily penalize profits. However, in such scenarios, Q_a approaches D_m, causing the denominator in the liquidity cost function to approach zero without penalizing profits sooner, which results in a steep drop in profit levels, making the curve less useful for explanatory purposes[10]. This illustrates a trade-off in the choice of parameters, which can be calibrated to reflect realistic market conditions, although such calibration is beyond the scope of the current paper.

[9]The maximizing value of Q_a given a D_m of 50 million is 8137103 . $8137103/50M = 0.1627$. From the balance sheet constrain we know $R = 50M - 8137103 = 41862897$. $41862897/50M = 0.8373$.

[10]See Appendix B for an example of the latter.

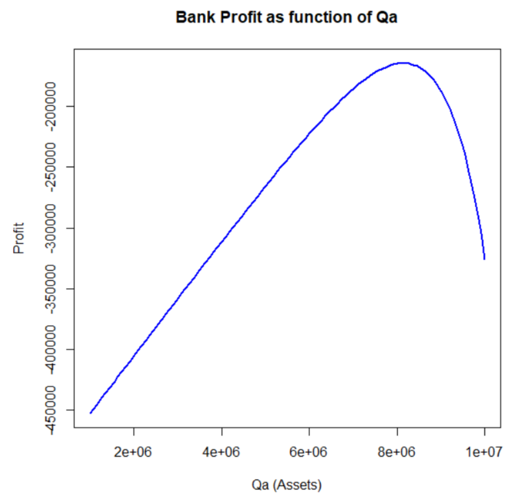

Figure 1: Bank profit Π as a function of loan quantity Q_a.

The plot shows the concave relationship between profit and Q_a, capturing the trade-off between interest income on loans, interest paid on demand deposits D_m, and increasing liquidity costs as reserves decline. The term $f(\theta)$ is omitted in this graph as it is unknown and does not affect the qualitative shape of the function. For the sake of plotting, the parameters are assumed as follows: $\alpha = 50$, $\gamma = 39$, $D_m = 50$ million, $i_l = 0.01$, and $i_a = 0.0475$. *Source:* Author's own elaboration using `R` software.

The foregoing shows how liquidty costs affect profit-maximizing behaviour. However, it may not be immediately obvious how this happens. The key mechanism lies in the heightened risk associated with diminished liquidity. In financial theory, risk is a cost: future income streams are discounted in proportion to their uncertainty. A competitive, profit-maximizing bank is therefore concerned not only with current margins but also with long-run sustainability. Increased risk from low reserve levels can manifest in several ways: depositors may become wary of the bank's financial health and withdraw funds, which could spur a redemption run; interbank counterparties may demand higher interest rates or deny credit; private clearinghouses may impose stricter liquidity requirements; and, critically, the bank may face solvency pressures if it cannot meet clearing obligations and is forced to liquidate assets at a loss. These potential costs shape the bank's strategic decisions, anchoring its behavior in the equilibrium condition derived from the first-order condition of profit maximization.

This interpretation is consistent with recent empirical research on the causes of bank failure. Correia, Luck, and Verner (2024) identify three robust historical patterns among failing banks over the past 160 years: (1) rising losses and worsening solvency precede failure; (2) failing banks increasingly substitute stable funding with non-core liabilities; and (3) these institutions exhibit a distinct boom-bust cycle, typically marked by rapid asset growth concentrated in illiquid loans. Similarly, Jamilov et al. (2024) find that deposit outflows are strongly predicted by bank leverage and weak profitability, highlighting how market participants respond preemptively to signs of vulnerability. These findings reinforce the theoretical insight that rising liquidity costs serve as a proxy for greater systemic risk, incentivizing banks to avoid aggressive balance sheet expansion when reserves are low.

It is equally important to recognize that a unilateral reduction in Q_a under conditions of constant demand for inside money implies that the bank is not maximizing profit, as its liquidity costs have not increased. The bank is not maximizing profit either if D_m increases and Q_a does not adjust accordingly. In such a scenario, the bank would be left with an unnecessarily conservative balance sheet, offering a lower return to liability holders than competitors might. Over time, this inefficiency would expose the bank to competitive pressures, as rival institutions offering a closer match between asset creation and liability demand could attract its depositors.

Again, this dynamic is particularly noteworthy: it suggests that profit-maximizing banks operating within a competitive free banking system have no inherent incentive to expand the supply of inside money (Q_l) unless it is in response to an increase in demand (D_m). In other words, the model implies that the supply of money (S_m) is endogenously determined by demand, such that $S_m = D_m$, aligning monetary expansion with market preferences. This mechanism provides theoretical support for the claim that free banking can maintain monetary equilibrium. In the following sections, we will discuss the potential implications, at a macroeconomic level, of monetary equilibrium.

3.4 Partial equilibrium model under aggregate D_m increase

A key question that emerges is: what happens when there is an aggregate increase in the demand for inside money? If interbank clearings perfectly offset one another, then no net change in reserves occurs, allowing banks to expand their loanable funds without constraint. Free banking theory offers a perspective on this scenario through the theory of precautionary reserves, which holds that as the scale of net interbank clearings grows, banks must maintain proportionally larger reserves to guard against payment imbalances.

Our model, with slight adjustments to its assumptions, can be adapted to explore this dynamic. First, we assume that reserves remain constant for any individual bank, reflecting the context in which aggregate demand for inside money increases but reserves do not shift between banks. As a result, reserves are no longer defined by the identity $R = Q_l - Q_a$. Additionally, we assume that under these conditions $Q_a = Q_l = D_m$, since any marginal increase in assets must be matched by an equal increase in liabilities. This leads to a modified profit function:

$$\Pi = f(\theta) + Q_a \cdot i_a - Q_l \cdot i_l - \alpha \left(\frac{Q_l}{R} \right)^{\gamma}$$

Substituting $Q_a = D_m$:

$$\Pi = f(\theta) + D_m \cdot i_a - D_m \cdot i_l - \alpha \left(\frac{D_m}{R} \right)^{\gamma}$$

Figure 2 illustrates how, under the assumption of constant reserves, our model captures the effect of an aggregate increase in the demand for inside money. Specifically, it shows that liquidity costs impose an upper bound on a bank's loan expansion. This outcome serves as a proxy of the *precautionary reserve theory*, as it constrains the growth of loans despite a continuous rise in the demand for money. Specifically, the law of precautionary reserve demand assumes that bank clearings rise or fall due to changes in frequency of payments. The total volume of clearings may also rise or fall because of an increase or decrease in the average size of individual payments where the frequency of payments is constant. This results in an increase in precautionary reserve demand proportional to the increase in bank clearings.

An important caveat is derived from the foregoing. As D_m increases the capacity of banks to expand their balance sheets declines. Potentially, this may constrain the creation of liabilities, such that $D_m < S_m$ — that is, the supply of inside money might fall short of demand. However, in a free banking environment with secular increases in productivity, long-run deflation may dampen nominal money demand growth. Whether the upper bound implied by the model ultimately limits the system's capacity to meet future demand is an empirical question, depending on the balance between monetary accommodation and the trajectory of nominal demand. This issue merits further investigation, which this model could provide the basis for.

Again, even under aggregate increases in the demand for inside money, profit-maximizing banks face a clear upper bound on loan creation. This reflects the inherent limitations imposed by liquidity costs on profitable expansion.

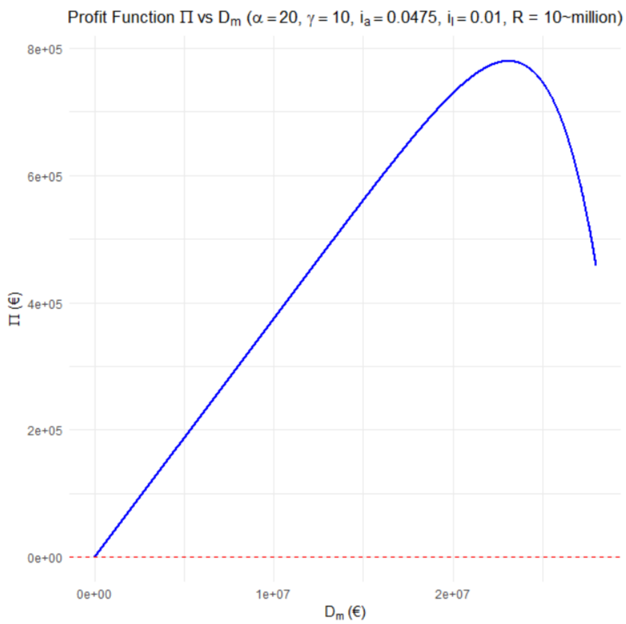

Figure 2: Banks profit Π as a function of Demand for inside-money D_m.

The plot shows the concave relationship between profit and D_m, capturing the trade-off between interest income on loans, interest paid on demand deposits, and increasing liquidity costs. The term $f(\theta)$ is omitted in this graph as it is unknown and does not affect the qualitative shape of the function. For the sake of plotting, the parameters are assumed as follows: $\alpha = 20$, $\gamma = 10$, $D_m = 50$ million, $i_l = 0.01$, $i_a = 0.06$ and $R = 10 Million$. *Source:* author's own elaboration using R software.

4 Macroeconomic implications of monetary equilibrium

At the macroeconomic level, the demand for inside money D_m is no longer treated as exogenous. Instead, it is now understood as an endogenous variable that reflects the public's aggregate liquidity preferences, shaped by broader economic conditions such as output, income, and expectations. This shift in perspective allows us to explore how systemic interactions among banks and the public give rise to a macro-level monetary equilibrium in which the supply of inside money adjusts to meet evolving demand.

4.1 Price level

To understand the price level implications of the previously established equilibrium condition $D_m = S_m$[11] , we take the classical quantity theory of money:

$$MV = PY$$

Solving for the price level:

$$P = \frac{MV}{Y}$$

This is the standard expression of the price level under the quantity theory of money. Since banks issue liabilities in response to profit opportunities—as derived in the model above—any autonomous changes in velocity (V) would be offset by endogenous adjustments in the money supply (M). For example, a decrease in V (reflecting higher demand for money balances) would incentivize banks to expand M through increased accommodation, while an increase in V would lead to contraction (Hendrickson 2019). Hence, the system tends to stabilize MV, allowing real output Y to determine the price

[11] This result relies on the assumption of smooth interbank coordination and rapid adjustment to changes in money demand. In practice, temporary disequilibria may arise due to coordination failures, asymmetric information, or delayed clearing responses. Nonetheless, the model shows that competitive pressures and liquidity constraints ultimately align supply with demand in equilibrium.

level P without monetary disturbances. Therefore, constant nominal spending (MV) can be assumed. As a result, an increase (decrease) in real output Y leads to a decrease (increase) in P. Thus, we reach the conclusion that under free banking with endogenous money supply determined by demand, the price level naturally adjusts to reflect changes in real output.

I acknowledge that the temporal dimension of the adjustment process, not taken into account in the model, is particularly relevant for assessing its macroeconomic implications. Although any unjustified increase in interest-earning assets (Q_a) is ultimately unsustainable due to liquidity constraints, if the adjustment occurs only after a significant delay, it may generate short-run distortions in prices or credit allocation. However, real-world evidence suggests that these adjustments occur rapidly. For example, deposits can be claimed by other banks almost instantaneously, as clearing and settlement operations are conducted frequently—often multiple times per day (Treasury 2021). Furthermore, in one of the most historically robust examples of a free banking system, Robert Somers (1873) noted that in 19th-century Scotland notes convertible on demand are constantly coming in for payment, and a large reserve of cash has to be kept and replenished in order to pay them. The average currency of a banknote in Scotland is ten or eleven days. This historical precedent reinforces the assumption that market discipline exerts strong and swift constraints on over-issuance.

In summary, when the demand for money (D_m) equals the supply of money (S_m), changes in the price level (P) are driven by changes in real output (Y). A negative supply shock, for instance, would lead to higher prices, while an increase in productivity would induce a decline in the price level. This outcome stands in contrast to the standard view, which treats a stable price index as the benchmark for monetary policy. From the perspective of monetary equilibrium theory, the price level is not anchored to a fixed nominal path—as would be the case under strict NGDP targeting—but instead serves as a flexible signal that conveys variations in productive efficiency. In this regard, our conclusions are in line with Salter and Young (2018), who argue that free banking systems adjust the money supply in response to changes in money demand, not in pursuit of a fixed nominal aggregate target. Thus, monetary stability in this context does not imply price-level stability, but rather the absence of monetary distortion in the transmission of real economic signals.

4.2 The natural rate of interest

An important implication of the identity $D_m = S_m$ is that an increase in the demand for bank liabilities constitutes an increase in voluntary savings. In a free banking system, this higher demand for inside money allows banks to expand their asset portfolios—specifically, by issuing new loans—without triggering disequilibrium, as shown in section 3. This process operates as follows: when agents choose to hold a greater quantity of bank-issued money, they are effectively postponing consumption, which reflects an intertemporal reallocation of resources. Banks, in turn, accommodate this increased demand for liabilities by creating offsetting interest-earning assets. As a result, the increased supply of loanable funds lowers interest rates, both on deposits and on bank loans, ensuring consistency between the monetary interest rate and the real intertemporal price of resources.

This mechanism ensures that the natural interest rate—i.e., the rate consistent with underlying time preferences or as defined by Wicksell the marginal productivity of capital —is reflected accurately in market outcomes. As Warburton (1950) points out, changes in the quantity of money which are not consonant with the rate of expansion needed for equilibrium also change the amount of funds available in the money loan market; thus they constitute the force which produces a departure of the market rate of interest from the equilibrium rate.

In this sense, equilibrium in the money market, defined as a match between the demand for and supply of inside money, is also an equilibrium in the market for loanable funds. Therefore, the workings of a competitive banking system align the supply of credit with the public's willingness to save, maintaining intertemporal coordination without central intervention. This highlights the inherent self-regulating nature of free banking regimes in aligning monetary and real variables.

4.3 Discussion of the macroeconomic implications

Sections 4.1 and 4.2 show that a banking system operating without a central bank is capable of equating money demand with money supply. Achieving monetary equilibrium tends to generate two key outcomes: first, the real interest rate converges to the natural rate; and second, changes in the price level are driven solely by productivity shocks. These results prompt a crucial normative question: is such an outcome desirable? In particular, what are the potential consequences of (1) deviations of the market interest rate from the natural rate, and (2) fluctuations in the price level that are not attributable to changes in productivity?

The core intuition here is that deviations of the real interest rate from its natural level, which may lead to demand-driven inflation or deflation, can be understood as two sides of the same coin[12] —especially when banks are viewed as intermediaries in the loanable funds market. In this role, banks simultaneously allocate capital and create inside money. When the market interest rate is set above the natural rate, the volume of credit creation is restricted. This reduces the quantity of inside money circulating in the economy, leading to a contraction in aggregate demand and downward pressure on prices—that is, deflation. In a world with sticky prices, that deflation can be economically painful, as described later on. At the same time, capital that could have been invested is withheld, further weakening economic activity. Conversely, if the interest rate falls below the natural rate, banks tend to expand credit beyond what is backed by voluntary savings. This excess issuance of inside money stimulates aggregate demand, placing upward pressure on prices and generating demand-driven inflation. According to Wicksell, if the market interest rate is below the natural interest rate, investment will exceed saving, leading to an increase in aggregate demand and consequently to inflation. On the other hand, if the market interest rate is above the natural interest rate, saving will exceed investment, resulting in a decrease in aggregate demand and potentially in deflation. These dynamics underscore the dual role of banks in shaping both the capital stock and the nominal expenditure path of the economy, thereby linking interest rate misalignment to monetary disequilibrium and price level instability.

An explanation of the potential consequences of deviations from the outcome produced by monetary equilibrium becomes pertinent. Firstly, the failure of a monetary system to ensure that the market real interest rate aligns with the natural rate (r^*) can lead to significant economic inefficiencies. When the real rate exceeds the natural rate, capital that could have been productively

[12] This is precisely the rationale behind the modern central bank practice of using short-term interest rate adjustments as the main tool for inflation control. By targeting the market interest rate to align with the estimated natural rate, central banks aim to stabilize aggregate demand and maintain price stability, thereby trying to avoiding both inflationary and deflationary spirals.

invested is withheld, suppressing aggregate demand and potentially inducing deflation. In such cases, the supply of money may fall short of agents' demand for real balances, exacerbating the downturn through a contraction in consumption and investment.

This mechanism is formally captured in the New Keynesian IS equation, which relates output to deviations of the real interest rate from its natural level. In its simplified form, the equation can be written as:

$$Y_t = Y^* - \alpha(r_t - r^*)$$

where Y_t denotes actual output, Y^* the economy's potential (or natural) output, r_t the real interest rate, and r^* the natural real interest rate. The parameter $\alpha > 0$ measures the sensitivity of output to interest rate deviations. This formulation illustrates that when the real rate r_t rises above its natural counterpart r^*, the output gap $(Y_t - Y^*)$ becomes negative—indicating a shortfall in aggregate demand. In essence, the equation captures how excessively tight monetary conditions, by elevating the real rate above its equilibrium, suppress economic activity relative to its long-run potential.

When the real interest rate r_t falls below the natural rate r^*, the IS equation likewise implies a positive output gap, as actual output Y_t exceeds its long-run potential Y^*. In this case, the term $(r_t - r^*)$ becomes negative, and the equation indicates that output rises above its sustainable trend. However, such expansions are not necessarily benign or efficient. Because New Keynesian models treat Y^* as the economy's long-run sustainable output level, any deviation above this level—i.e., a positive output gap—is implicitly assumed to be temporary. Such a gap may signal overheating, leading to inflationary pressures and prompting monetary tightening. The return to potential output often involves an economic slowdown, and in some cases, a painful downturn. Thus, even from a mainstream perspective, sustained output above potential is not viewed as benign. When combined with the Austrian insight that artificially low interest rates may lead to distorted investment patterns (Mises 1949), the risks of such booms become even more pronounced. These expansions may not only be unsustainable but also embed structural imbalances that manifest during the subsequent correction.

A further concern is the potential endogeneity of the natural rate itself. A study by the Centre for Economic Policy Research (CEPR) highlights that persistently low real interest rates can reduce the natural rate over time by sustaining unproductive firms and depressing total factor productivity (TFP) (End and Hoeberichts 2018).

Symmetrically, a rise in the real interest rate may contribute to a future increase in the natural rate by reallocating resources toward more productive uses. However, this process may come at the cost of lower present consumption, particularly through higher borrowing costs such as mortgage rates.

As for the price level, demand-driven inflation—defined as a general rise in prices due to excessive growth in nominal spending not matched by real output—can have several distortionary effects on economic activity and income distribution[13]. First, according to the Cantillon effect, the injection of new money into the economy does not affect all agents simultaneously or proportionally. Those who receive the new money earlier, typically financial institutions and well-capitalized investors, benefit from increased purchasing power before prices have fully adjusted. Conversely, wage earners and fixed-income households experience a decline in real income, as their nominal earnings lag behind rising prices and a inflation tax effect. In a recent study Charalampakis et al. (2022), show how this redistribution effect disproportionately harms the poor, who have fewer opportunities to hedge against inflation or access inflation-protected assets[14]. Second, persistent inflation increases uncertainty about the future price level, making long-term planning, investment, and contract negotiation more difficult. This heightened uncertainty can deter productive investment and shorten planning horizons, thereby reducing long-term growth. Finally, inflation can distort relative prices, leading to misallocation of resources (Hayek 1976). Firms may mistake general price increases for sector-specific demand, allocating capital in-

[13] The mechanisms described assume that agents do not possess fully rational expectations in the Lucasian sense. Under rational expectations, some distortions—such as misallocation due to perceived relative price changes or delayed consumption decisions—may be mitigated. However, even under forward-looking behavior, incomplete information, nominal rigidities, and financial frictions can still result in significant real effects from demand-driven inflation or deflation (Lucas 1995).

[14] The interpretation of the inflation period as demand driven shock is, at least partially, supported by recent empirical research. Di Giovanni et al. (2023) demonstrate that a significant share of the inflation observed during the pandemic period in Europe can be attributed to demand-side factors rather than purely supply constraints. Their multi-country New Keynesian model indicates that the fiscal and monetary responses to the pandemic played a notable role in stimulating aggregate demand, thereby contributing to inflationary pressures.

efficiently and contributing to unsustainable production patterns.

On the other end of the spectrum, demand-driven deflation—typically the result of insufficient monetary expansion or excessively tight credit conditions—also poses serious risks to macroeconomic stability. A generalized fall in prices reduces current consumption, as households postpone purchases in anticipation of lower future prices. This hoarding behavior depresses aggregate demand, deepening economic contractions. Deflation also increases the real burden of nominal debts. Firms and households that borrowed under previous price expectations find it increasingly difficult to meet fixed repayment schedules, leading to higher default rates, banking sector stress, and potential financial instability. Furthermore, falling prices compress firms' profit margins, prompting cost-cutting, wage reductions, or layoffs. The resulting decline in income further suppresses demand, creating a deflationary spiral. Unlike supply-driven deflation (associated with productivity gains), demand-driven deflation is typically harmful, as it reflects a failure of the monetary system to accommodate agents' desire to hold liquid balances, rather than an improvement in real economic fundamentals.

This reasoning aligns with a wide array of macroeconomic frameworks. Economists like the mentioned Knut Wicksell and modern advocates of the natural interest rate concept argue for the importance of aligning policy rates with r^* to avoid inflationary or deflationary imbalances.

While the preceding analysis underscores the theoretical appeal of a free banking system in aligning the market real interest rate with the natural rate (r^*) and price movements to output shocks, it is essential to recognize that this alignment may not always yield optimal macroeconomic outcomes. Particularly, when considerations such as full employment and liquidity traps are introduced, alternative policy interventions might become pertinent.

From a Keynesian perspective, moderate inflation can play a crucial role in addressing wage stickiness—a phenomenon where nominal wages are slow to adjust downward due to institutional contracts, worker morale, and minimum wage laws. In such contexts, inflation effectively reduces real wages, thereby enhancing labor demand and facilitating a return to full employment.

Moreover, scenarios characterized by liquidity traps can be an issue. In a free banking framework, an analog to the Keynesian liquidity trap might emerge when the demand for bank-issued liabilities becomes infinitely elastic with respect to interest rates. In such a case, even if banks extend issuance and offer negative interest rates on deposits, agents may prefer to hold onto their balances rather than spend, rendering monetary expansion ineffective. This could result in a self-reinforcing contraction of aggregate demand and deflationary pressures. While this scenario is theoretically consistent with liquidity trap logic, its empirical relevance under historical free banking regimes remains uncertain (Selgin 1988b). In such a case Keynesians advocate for expansionary fiscal policy. Here, increased government spending directly boosts aggregate demand, circumventing the initial liquidity trap.

Conversely, the Pigou effect posits that deflation increases the real value of money holdings, thereby stimulating consumption and, subsequently, employment (Mansoorian 2012). This mechanism suggests that, under certain conditions, demand-driven deflation can self-correct economic downturns. While theoretically plausible, the empirical relevance of the Pigou effect is widely debated. Historical episodes such as the Great Depression illustrate that falling prices did not lead to increased real balances and higher consumption, but rather coincided with steep declines in output and persistent unemployment. Friedman and Friedman (1980) argue that the Federal Reserve's failure to prevent a collapse in the money supply was the principal cause of the Depression, challenging the notion that deflation can self-correct via real balance effects[15].

Furthermore, societal preferences may sometimes favor short-term economic relief over long-term efficiency. For instance, in the face of severe recessions, policymakers might deliberately lower real interest rates below the natural rate, accepting the trade-off of future inflation (and other potential consequences) to mitigate immediate economic pain.

These considerations imply that, despite its theoretical merits, a free banking system's outcomes may not always align with broader macroeconomic objectives, necessitating a further nuanced discussion.

[15]While monetary explanations are crucial, it is equally important to consider non-monetary mechanisms, as emphasized by The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (2022). Bernanke's research highlights how disruptions in financial intermediation—particularly constraints on bank credit—intensified the economic downturn during the Great Depression, compounding the effects of falling money supply.

5 Conclusion

This paper has developed a formal, micro-founded model of bank behavior under a free banking regime, with the aim of investigating whether such a system can yield monetary equilibrium through decentralized, profit-maximizing actions. The core insight is that when a bank faces increasing liquidity costs as reserves decline, it will endogenously limit its issuance of interest-earning assets. This generates an interior profit-maximizing level of loan creation (Q_a), and implies that the supply of inside money (S_m) adjusts to match the demand for it (D_m), such that $S_m = D_m$. This equilibrium constraint emerges not from regulation, but from the internal logic of bank optimization under competition.

At the macroeconomic level, this self-regulating mechanism has important implications. When money supply tracks demand, the price level becomes a function of real output through the classical quantity equation $MV = PY$. Under free banking, the endogenous expansion or contraction of inside money offsets changes in velocity, maintaining nominal spending stability. Consequently, the price level reflects real productivity changes, not monetary distortions. Similarly, because increased money demand reflects greater voluntary saving, the system naturally accommodates this with loanable funds expansion, leading to an interest rate that converges to the natural rate. These dynamics offer a decentralized route to macroeconomic coordination—aligning credit supply with real intertemporal preferences.

However, while these outcomes are promising, they do not necessarily imply optimality under all conditions. As discussed, under nominal rigidities or liquidity traps, deflationary pressures—even when consistent with equilibrium—may generate real economic costs. Moreover, the model abstracts from frictions like asymmetric information, temporal lags in adjustment, and institutional complexities. Therefore, although the results support the internal consistency of free banking theory, caution is warranted in extrapolating normative conclusions.

From a policy perspective, this work contributes to the ongoing reassessment of the institutional foundations of money and banking. While it does not challenge the logical assumptions of free banking theory—those are taken as given—it offers a formal framework for understanding their implications. From a Bayesian standpoint, the exercise of formalization itself lends incremental credibility to free banking as a viable alternative, potentially nudging the prior beliefs of economists and policymakers. If such theoretical work contributes to a shift in beliefs—even marginally—it becomes highly relevant, as reconfiguring the monetary regime would have profound political and institutional consequences.

This model also lays the groundwork for further inquiry. Empirically, it raises the question of whether historical or hypothetical free banking systems can generate sufficient S_m to meet rising D_m, especially under secular deflation. The model could be extended to explore whether nominal demand shortfalls arise from the upper bounds imposed by liquidity costs. Additionally, future iterations should consider non-continuous or stochastic liquidity costs to better reflect real-world banking crises and nonlinear responses to risk. A particularly important enhancement would be to model the interaction between the natural interest rate and the bank's ability to supply Q_a—a relationship currently assumed fixed in this framework and that should be endogenized. Another area for improvement of the current model is the absence of default risk. In reality, bank lending entails uncertainty over repayment, and the profitability of extending credit is influenced not only by liquidity constraints but also by the expected return distribution of the underlying assets. A valuable extension would be to introduce stochastic default probabilities into the model, making the bank's expected return on Q_a a function of both interest income and credit risk.

In sum, this paper offers a tractable and internally coherent model that strengthens the case for the plausibility of free banking systems to maintain monetary equilibrium. It neither idealizes nor dismisses such regimes, but rather seeks to understand their internal mechanics. By doing so, it contributes to the broader debate on monetary institutions and invites further formal and empirical exploration.

References

Alesina, Alberto (1988). "Macroeconomics and Politics". In: *NBER Macroeconomics Annual 1988, Volume 3*. Ed. by Stanley Fischer. MIT Press, pp. 13–62. URL: http://www.nber.org/chapters/c10951.

Anginer, Deniz and Asli Demirgüç-Kunt (2018). *Bank Runs and Moral Hazard: A Review of Deposit Insurance*. Policy Research Working Paper 8589. The World Bank. DOI: 10.1596/1813-9450-8589. URL: https://openknowledge.worldbank.org/handle/10986/30398.

Azariadis, Costas and Leo Kaas (2023). "Private and Public Circulating Liabilities". In: *Review of Economic Dynamics* 49, pp. 68–90. DOI: 10.1016/j.red.2022.10.006.

Cavalcanti, Ricardo de O., Andrés Erosa, and Ted Temzelides (1999). "Private Money and Reserve Management in a Random-Matching Model". In: *Journal of Political Economy* 107.5, pp. 929–945. DOI: 10.1086/250085.

Charalampakis, Evangelos et al. (2022). "The impact of the recent rise in inflation on low-income households". In: *ECB Economic Bulletin* Issue 7. URL: https://www.ecb.europa.eu/press/economic-bulletin/focus/2022/html/ecb.ebbox202207_04~a89ec1a6fe.en.html.

Correia, Sergio A., Stephan Luck, and Emil Verner (2024). *Failing Banks*. Tech. rep. 32907. National Bureau of Economic Research. URL: https://www.nber.org/papers/w32907.

Diamond, Douglas W. and Philip H. Dybvig (1983). "Bank Runs, Deposit Insurance, and Liquidity". In: *Journal of Political Economy* 91.3, pp. 401–419. DOI: 10.1086/261155.

Dowd, Kevin (1996). "The Case for Financial Laissez-Faire". In: *The Economic Journal* 106.436, pp. 679–687. DOI: 10.2307/2235576. URL: https://doi.org/10.2307/2235576.

End, Jan Willem van den and Marco Hoeberichts (2018). *The Curse of Persistently Low Real Interest Rates*. https://cepr.org/voxeu/columns/curse-persistently-low-real-interest-rates. VoxEU Column, Centre for Economic Policy Research, 25 April 2018.

Federal Financial Institutions Examination Council (US) and Federal Reserve Bank of St. Louis (2020). *Net Interest Margin for all U.S. Banks (DISCONTINUED)*. https://fred.stlouisfed.org/series/USNIM. Accessed: June 19, 2025.

Friedman, Milton and Rose Friedman (1980). *Free to Choose: A Personal Statement*. New York: Harcourt.

Hayek, Friedrich A. (1976). *Denationalisation of Money: The Argument Refined*. 1st. An analysis of the implications of abolishing government monopoly on money issuance. London: Institute of Economic Affairs.

Hendrickson, Joshua R. (2019). *How Should the Central Bank Determine Its Target for Nominal Income Growth? Lessons from Free Banking and

Friedman. Policy Brief. Accessed: 2025-05-19. Mercatus Center at George Mason University. URL: https://www.mercatus.org/research/policy-briefs/how-should-central-bank-determine-its-target-nominal-income-growth.

Jamilov, Rustam et al. (2024). *Two Centuries of Systemic Bank Runs*. Department of Economics Discussion Paper 1039. University of Oxford. URL: https://ora.ouls.ox.ac.uk/objects/uuid:1f83d78d-4282-4d94-9c78-2f6e57362c2f.

Lucas, Robert E. (1995). *Nobel Lecture: Monetary Neutrality*. https://www.nobelprize.org/prizes/economic-sciences/1995/lucas/lecture/. Nobel Prize Lecture, December 7, 1995.

Mansoorian, Arman (2012). "Money, capital and the Pigou effect". In: *Economics Letters* 117.1, pp. 354–356. DOI: 10.1016/j.econlet.2012.05.028. URL: https://www.sciencedirect.com/science/article/pii/S0165176512002960.

McLeay, Michael, Amar Radia, and Ryland Thomas (2014). "Money creation in the modern economy". In: *Bank of England Quarterly Bulletin* 54.1. Bank of England, pp. 14–27. URL: https://www.bankofengland.co.uk/-/media/boe/files/quarterly-bulletin/2014/money-creation-in-the-modern-economy.

Mises, Ludwig von (1949). *Human Action: A Treatise on Economics*. New Haven: Yale University Press.

Salter, Alexander William and Andrew T. Young (2018). "Would a Free Banking System Stabilize NGDP Growth?" In: *The Quarterly Review of Economics and Finance* 70. Forthcoming; preprint posted on SSRN April 24, 2018, pp. 21–25. DOI: 10.1016/j.qref.2018.04.016.

Sanches, Daniel R. (2016). "On the inherent instability of private money". In: *Review of Economic Dynamics* 21, pp. 126–144. DOI: 10.1016/j.red.2015.06.001.

Selgin, George A. (1988a). "In Defence of Fiduciary Media". In: *Economic Journal* 98.392, pp. 443–463.

– (1988b). *The Theory of Free Banking: Money Supply under Competitive Note Issue*. Totowa, NJ: Rowman & Littlefield.

Somers, Robert (1873). *The Scotch Banks and System of Issue: Including Translation of Les Banques d'Écosse*. Edinburgh: MacLaren and Macniven.

The Committee for the Prize in Economic Sciences in Memory of Alfred Nobel (2022). *Financial Intermediation and the Economy*. Scientific background on the Sveriges Riksbank Prize in Economic Sciences in Memory of

Alfred Nobel 2022. URL: https://www.nobelprize.org/uploads/2022/10/advanced-economicsciencesprize2022.pdf.
Treasury, Modern (2021). *The Difference Between Settlement and Clearing and Why it Matters to Banks*. Accessed: May 12, 2025. URL: https://www.moderntreasury.com/journal/difference-between-settlement-and-clearing.
Warburton, Clark (1950). "The Monetary Disequilibrium Hypothesis". In: *The American Journal of Economics and Sociology* 10.1, pp. 1–11. URL: https://www.jstor.org/stable/3483424.

Appendices

A Derivation of the Effect of D_m on Q_a

We start from the profit function:

$$\Pi = Q_a \cdot i_a - D_m \cdot i_l - \alpha \left(\frac{D_m}{D_m - Q_a} \right)^{\gamma}$$

Partial Derivative:

$$\frac{\partial \Pi}{\partial Q_a} = i_a - \alpha \cdot \gamma \cdot \left(\frac{D_m}{D_m - Q_a} \right)^{\gamma - 1} \cdot \frac{D_m}{(D_m - Q_a)^2}$$

Solving for Q_a, we obtain:

$$Q_a = D_m - \left(\frac{\alpha \cdot \gamma \cdot D_m^{\gamma}}{i_a} \right)^{\frac{1}{\gamma+1}}$$

Define:

$$A = \left(\frac{\alpha \cdot \gamma \cdot D_m^{\gamma}}{i_a} \right)^{\frac{1}{\gamma+1}} \quad \Rightarrow \quad Q_a = D_m - A$$

Differentiate with respect to D_m:

$$\frac{dQ_a}{dD_m} = 1 - \frac{dA}{dD_m}$$

Where:

$$\frac{dA}{dD_m} = \frac{1}{\gamma+1} \cdot \left(\frac{\alpha \cdot \gamma}{i_a}\right)^{\frac{1}{\gamma+1}} \cdot \gamma \cdot D_m^{\frac{\gamma-1}{\gamma+1}} > 0$$

Since all parameters are positive, $\frac{dA}{dD_m}$ is strictly positive and increasing in D_m.

Therefore:

$$\frac{dQ_a}{dD_m} = 1 - \frac{dA}{dD_m}$$

is initially positive, but eventually becomes negative as D_m increases and $\frac{dA}{dD_m} > 1$. This implies a non-monotonic relationship between D_m and Q_a.

We now show that $Q_a(D_m)$ is concave by computing the second derivative. Recall the expression:

$$Q_a = D_m - \left(\frac{\alpha\gamma D_m^\gamma}{i_a}\right)^{\frac{1}{\gamma+1}}$$

Define:

$$A = \left(\frac{\alpha\gamma D_m^\gamma}{i_a}\right)^{\frac{1}{\gamma+1}} = C \cdot D_m^{\frac{\gamma}{\gamma+1}}, \quad \text{where} \quad C = \left(\frac{\alpha\gamma}{i_a}\right)^{\frac{1}{\gamma+1}}$$

Then:

$$\frac{dQ_a}{dD_m} = 1 - \frac{dA}{dD_m} = 1 - C \cdot \frac{\gamma}{\gamma+1} \cdot D_m^{\frac{\gamma-1}{\gamma+1}}$$

and:

$$\frac{d^2Q_a}{dD_m^2} = -\frac{d^2A}{dD_m^2} = -C \cdot \frac{\gamma}{\gamma+1} \cdot \frac{\gamma-1}{\gamma+1} \cdot D_m^{\frac{\gamma-2}{\gamma+1}} < 0$$

Since all parameters are positive and $\gamma > 2$, it follows that $\frac{d^2Q_a}{dD_m^2} < 0$. Therefore, $Q_a(D_m)$ is strictly concave.

B Bank profit as a function of Q_a

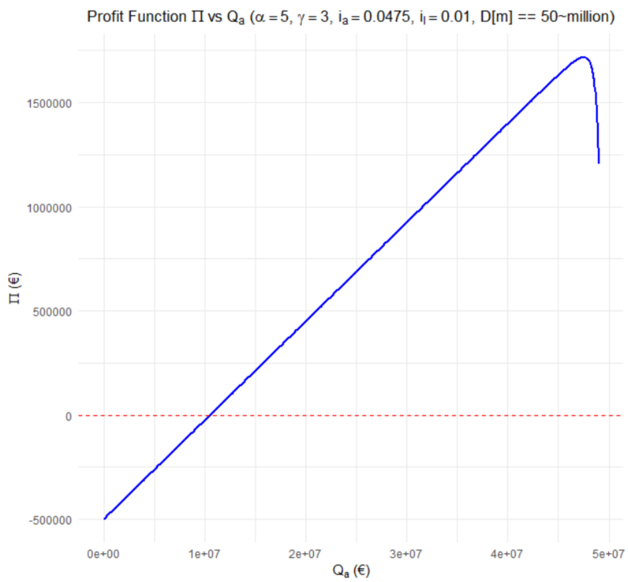

Figure 3: Banks profit Π as a function of loan quantity Q_a (Different parameters).

The plot shows the concave relationship between profit and Q_a, capturing the trade-off between interest income on loans, interest paid on demand deposits D_m, and increasing liquidity costs as reserves decline. The term $f(\theta)$ is omitted in this graph as it is unknown and does not affect the qualitative shape of the function. For the sake of plotting, the parameters are assumed as follows: $\alpha = 5$, $\gamma = 3$, $D_m = 50$ million, $i_l = 0.01$, and $i_a = 0.0475$. The profit maximizing Q_a in this case is 47.48 Million. Which makes up to 94% of D_m. $47.48/50 = 0.9496$, meaning that Reserves make up to close to 5% of total liabilities (D_m). This is more in line with real world figures.

Source: author's own elaboration using `R` software.