University of Zurich[UZH]

University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 310

A Model of Endogenous Financial Inclusion: Implications for Inequality and Monetary Policy

Mohammed Ait Lahcen and Pedro Gomis-Porqueras

Revised version, February 2021

A Model of Endogenous Financial Inclusion: Implications for Inequality and Monetary Policy*

Mohammed Ait Lahcen† Pedro Gomis-Porqueras‡

Qatar University Deakin University

This Version: February 17, 2021

Abstract

We propose a monetary model with endogenous credit market participation to study the impact of financial inclusion on inequality and welfare. We find that consumption inequality results from differences in agents' decision to access financial services. This heterogeneity generates a pecuniary externality, potentially resulting in some agents over-consuming. Moreover, monetary policy has distributional consequences. To quantify these effects, we calibrate our model to India, accounting for a third of observed consumption inequality. Finally, we analyze various policies aimed at increasing financial inclusion and find that a direct transfer to bank account holders yields the highest welfare and lowest consumption inequality.

Keywords: money; credit; banking; financial inclusion; inequality.

JEL classification: E40, E50.

---

*We would like to thank the Editor and two anonymous referees for their valuable feedback and suggestions. We are also grateful to Lukas Altermatt, Aleksander Berentsen, Gabriele Camera, Jonathan Chiu, Cyril Monnet and Guillaume Rocheteau for their valuable feedback. We thank seminar participants at the Universities of Basel and Bern, the 2018 International Days of Macroeconomics and Finance at the Central Bank of Morocco, the 2018 Australasian Meeting of the Econometric Society at Auckland University of Technology, the 2018 St. Louis Fed Summer Workshop for Money, Banking, Payments and Finance, the 2018 Annual Congress of the European Economic Association at the University of Cologne, the 2018 Meeting of the German Economic Association at the University of Freiburg and the 2018 Gerzensee Alumni conference for valuable comments and suggestions. Ait Lahcen would like to acknowledge the financial support of the Swiss National Science Foundation and the Freiwillige Akademische Gesellschaft Basel.

†Department of Finance and Economics, Qatar University, P. O. Box: 2713, Doha, Qatar. e-mail: m.aitlahcen@gmail.com

‡Department of Economics, Deakin University, Burwood, Vic 3145, Australia. e-mail: peregomis@gmail.com

1 Introduction

Financial exclusion (the lack of access to basic financial services) is a widely observed phenomenon in developing countries. According to Demirgüç-Kunt et al. (2015), 94% of the adult population in OECD countries owns an account at a formal financial institution. This proportion, however, is only 54% in developing nations, ranging from 14% in the Middle East and North Africa to 69% in East Asia and the Pacific. This type of market segmentation has important implications for the rest of the economy. For instance, using aggregate macroeconomic data, there exists a robust negative correlation between different measures of financial inclusion and consumption inequality across countries.[1] In addition, many microeconomic studies suggest that improved access to finance reduces income inequality, poverty and increases food security.[2] Moreover, some authors argue that increasing financial inclusion can affect the impact and effectiveness of monetary policy. This is the case as wider access to saving vehicles makes consumers more reactive to changes in interest rates, improving the transmission of monetary policy.[3] These different findings highlight the inherent relationship between access to financial markets, inequality and monetary policy. This paper studies such links.

To do so, we consider a monetary framework with endogenous credit market participation, where financial inclusion is an equilibrium outcome. Agents face idiosyncratic preference shocks that determine their willingness to consume in frictional markets. Since agents are anonymous when trading in frictional markets, fiat money becomes an essential medium of exchange. Thus, these shocks generate uncertainty about future liquidity needs. Agents have also access to a frictionless competitive market where they can produce and trade the numeraire good, rebalance their portfolios and settle their financial obligations. Consumers can borrow or deposit their money holdings by accessing a competitive banking sector. However, accessing banks is costly. This cost differs across consumers, capturing the different physical and informational costs that agents face when accessing banking services.[4] The measure of consumers who decide to incur these costs determines the degree of financial inclusion.

Within this environment, we find that differences in the cost of access to basic financial

[1] Using data from 159 countries, we find a negative correlation between financial inclusion, measured by account ownership, and consumption inequality. The magnitude of the correlation is about -0.4.

[2] We refer to Burgess and Pande (2005); Levine (2005); Beck et al. (2005, 2007); Marshall (2004); Sarma and Pais (2011); Laha et al. (2011) for more on these issues.

[3] See Mehrotra and Yetman (2014, 2015) for a related discussion.

[4] According to Allen et al. (2016) and Demirgüç-Kunt et al. (2015), lower financial inclusion is associated with higher fees, higher physical costs of accessing financial intermediaries, lack of necessary documentation, weaker legal rights and less political stability.

services can result in significant consumption inequality. Furthermore, the measure of financially included agents is non-monotonic in inflation and liquidity risk. Given that financially included and excluded agents coexist, monetary policy can have distributional consequences. This is the case as agents face different liquidity constraints and monetary policy affects them differentially. Moreover, under competitive pricing, this heterogeneity generates a pecuniary externality which can result in over-consumption of financially included agents compared to the socially efficient level. To determine the empirical relevance of our theoretical predictions, we conduct a quantitative assessment for the case of India. Our simple model is able to account for approximately a third of the observed consumption inequality. It accounts also for half of the share of consumer credit to GDP and 70% of the demand deposits to M1 ratio. Moreover, we show that recent changes in the costs of accessing banking services in India can account for more than a third of the observed increase in financial inclusion. Finally, we analyze various policies aimed at increasing financial inclusion. Because of the pecuniary externality, interest rate policies can result in a decrease in welfare and an increase in consumption inequality. We show that a borrowing interest rate subsidy is more distorting and costly than a deposit interest rate subsidy. Moreover, we find that a direct benefit transfer scheme, aimed at bank account owners, is preferred to interest rate policies. Such policy can reduce consumption inequality and increase welfare. This is the case even as individual consumption decreases. In light of these results and compared to traditional policy recommendations, this paper highlights the importance of providing adequate returns on deposits and direct benefit transfer schemes to bank users in increasing financial inclusion. It also brings attention to possible unintended consequences of higher access to banks on the remaining unbanked population.

2 Related literature

This paper connects with three different strands of the literature: that which explores the implications of limited access to financial markets, that which studies the coexistence of money and credit and finally that which explores inequality and inflation.

The seminal papers of Chatterjee and Corbae (1992), Allen and Gale (1994) and Williamson (1994) study the different type of equilibria that emerge when financial markets are endogenously segmented. In an environment where the demand for money results from transaction costs in other assets, Chatterjee and Corbae (1992) show that there may be an equity-efficiency trade-off stemming from monetary deflation. Allen and Gale (1994), in a Diamond and Dybvig (1983) environment, find that allowing for endogenous market participation can cause significant price volatility and generate multiple equilibria. In a similar vein, when an illiquid asset is subject to fixed transactions costs, Williamson (1994) shows that there exists a participation externality that tends to deliver under-provision of liquidity.[5] Relative to this literature, we study how endogenous financial participation is affected by monetary policy. We also analyze various policies aimed at increasing financial participation and study their implications for welfare and consumption inequality.

Frictions are necessary to generate an essential role for money as a medium of exchange.[6] The monetary literature has emphasized that some of these frictions prevent the use of alternative payment instruments.[7] Some authors, such as Rojas Breu (2013) and Chiu et al. (2018), have emphasized the importance of considering segmented credit markets. Our paper relates to this strand of the literature. Rojas Breu (2013) focuses on costless credit in an environment where uncertainty regarding access of agents to credit generates a precautionary demand for money. Chiu et al. (2018) focus on costly credit in an environment with exogenous limited credit market participation. Both of these papers show that an increase in credit market participation has an ambiguous effect on welfare. We contribute to this literature by endogenizing the participation margin. As a result, the occurrence of the pecuniary externality is not limited to exogenous changes in credit market participation. It can also result from any policy that affects agents' decision to participate in credit markets.

Finally, this paper relates to the monetary literature that studies the relationship between inflation and inequality. Countries with more unequal income distributions tend to have higher inflation.[8] There have been few attempts to rationalize this fact. Erosa and Ventura (2002) build a monetary growth model consistent with key features of cross-sectional household data and use this framework to study the distributional impact of inflation. The authors find that if credit services exhibit economies of scale, inflation can work as a non-linear regressive consumption tax. Gomis-Porqueras (2001), on the other hand, considers a monetary growth model where agents face a minimum bank deposit. He shows that limiting access of low-skilled workers to the banking sector results in lower inflation and a decrease in the investment of physical capital, increasing welfare and reducing wage inequality. Cysne et al. (2005), instead, consider a shopping-time framework where agents have different productivity levels and differentiated access to financial asset markets. The authors

[5]A similar result is highlighted by Berentsen et al. (2014).

[6]Absence of double coincidence of wants, spatial separation, absence of a record keeping technology and no commitment have been advocated to explain the use of money to facilitate exchange.

[7]Monnet and Roberds (2008), Bencivenga and Camera (2011),Sanches and Williamson (2010), Chiu and Meh (2011), Sanches (2011), Rojas Breu (2013), Lotz and Zhang (2016), Gu et al. (2016), Chiu et al. (2018), among others, study the coexistence of money and credit. See Lagos et al. (2017) for a recent review of the New Monetarist literature and Rocheteau and Nosal (2017) for a textbook treatment.

[8]We refer to Beetsma and Van Der Ploeg (1996); Romer and Romer (1999); Easterly and Fischer (2001); Albanesi (2007) among others, for more details about such findings.

show that there exists a positive correlation between inflation and income inequality. In contrast, Menna and Tirelli (2017) show, in a model with exogenous limited asset market participation, that higher inflation can reduce wealth inequality as it taxes the larger money holdings (in absolute terms) of wealthier individuals. In contrast to this literature, we focus on consumption inequality that results endogenously from differences in the cost of financial participation. We also show the existence of a positive relationship between this type of consumption inequality and inflation.

3 Environment

The general environment is based on Lagos and Wright (2005), Rocheteau and Wright (2005) and Berentsen, Camera, and Waller (2007). Time is discrete and continues forever. The economy is populated by two types of infinitely lived agents each of unit measure: buyers and sellers. Private agents trade in sequential goods markets that differ in terms of their frictions. In addition, agents have access to financial intermediaries operating in a competitive market to finance part of their consumption. Buyers and sellers discount the future at rate $\beta \in (0, 1)$.

Each period is divided into three consecutive sub-periods. In the first sub-period, buyers have access to a competitive banking sector, which offers loans and deposits. We call this market BM. In the second sub-period, buyers and sellers trade a specialized perishable good in an anonymous competitive market, which we refer to as AM. Finally, in the third sub-period, agents have access to a frictionless competitive market where they can produce and trade the numeraire perishable good, rebalance their portfolios and settle their financial obligations. We refer to this last market as CM.

Since buyers are anonymous and sellers do not have access to record-keeping services when trading in the AM, a medium of exchange is essential for trades to take place. In contrast, since agents can produce and consume the CM numeraire good, a medium of exchange in this market is not essential. The only durable asset in this economy is an intrinsically useless object issued by the government, namely fiat money, which we denote by M_t. The supply of money grows at a rate $\gamma > \beta$ and is injected (withdrawn) through lump sum transfers (taxes) in the CM.

Preferences and technologies: Buyers are subject to an idiosyncratic preference shock that affects their marginal utility of consumption in the AM. In particular, with probability σ, a buyer gets utility $u(q)$ of consuming q units of AM goods, where $u'(q) > 0$, $u''(q) < 0$, $u'(0) = +\infty$ and $u'(+\infty) = 0$. With probability $1 - \sigma$, a buyer obtains no utility from consuming the AM good. This preference shock is independent across buyers and time, and generates heterogeneity among buyers in terms of their liquidity needs.[9]

In the CM, all agents can consume and produce the CM good. By consuming x units of the CM good, the buyer obtains utility $U(x)$, where $U'(x) > 0$, $U''(x) \leq 0$, $U'(0) = +\infty$ and $U'(+\infty) = 0$. Agents incur a linear disutility when producing the CM good. Thus, the period utility of a buyer is given by

$$\mathcal{U}^b = \sigma u(q) + U(x) - x. \tag{1}$$

Sellers incur disutility $c(q)$ when producing q units of the AM good; where $c'(q) > 0$ and $c''(q) \geq 0$. Sellers can also produce x units of the numeraire good with linear disutility and consume it to get utility $U(x)$. Hence, sellers' period utility is given by

$$\mathcal{U}^s = -c(q) + U(x) - x. \tag{2}$$

As in Berentsen et al. (2007), financial intermediaries accept one-period nominal deposits and offer one-period nominal loans. From now on we refer to these intermediaries as banks. This is the case as they perform some of the banks' functions. Banks have access to a costless record keeping technology that allows them to register the identity of agents. Moreover, the government is able to enforce deposit and loan contracts in CM. These two assumptions make financial intermediation possible. We rule out issues of commitment and assume borrowers do not default on their loans and banks are fully committed to repaying their depositors.

Each buyer faces an idiosyncratic, fixed and time-invariant (permanent) cost ε that he must incur at the beginning of the CM in order to access banking services in the following period. This cost ε is distributed across buyers according to a continuous distribution $F(\varepsilon)$ with support $[\underline{\varepsilon}, \bar{\varepsilon}]$.[10] It is intended to capture the physical and informational costs that buyers face when accessing financial services.

Timing: Given their ε cost, buyers decide at the beginning of the CM whether to incur the cost and access the BM next period or not. At the beginning of the next period, buyers are subject to the preference shock. Once the shock is realized, buyers who have decided to access the BM can deposit or borrow money. After this market closes, all buyers and sellers trade in the AM and subsequently in the CM. To simplify notation, we drop the time index for current period variables and index the next and previous periods' variables by +1 and

[9] This setup is isomorphic to a model with decentralized trades and search frictions where the probability of finding a seller in the AM is σ.

[10] One way of thinking about this structure is to have every buyer always being in a specific location. The transport cost of visiting the closest bank from that location is ε. Some buyers live in remote villages and hence face a high ε, while others live in big cities, hence facing low ε.

-1 respectively.

4 Decentralized solution

In what follows we describe agents' decision problems and determine the resulting equilibria. We focus on stationary monetary equilibria where aggregate real balances are constant over time. This implies that $\phi M = \phi_{+1} M_{+1}$; where ϕ is the value of money in units of the CM good. Given the sequential nature of agents' decisions, we first start with the CM problem. We then separately analyze the optimal trading decisions in the AM of financially included and excluded agents. Finally, we solve the banks' problem.

4.1 Sellers' problem

CM problem: In order to focus our analysis on buyers, we assume that sellers do not have access to banking services and hence solve the following optimization problem[11]

$$W^s(m) = \max_{x,h,m_{+1}} U(x) - h + \beta V^s_{+1}(m_{+1}) \tag{3}$$

$$\text{s.t. } x + \phi m_{+1} = h + \phi m + \mathcal{T} \tag{4}$$

where V^s and W^s denote the AM and CM value functions, respectively, $\mathcal{T}$ is the real monetary lump sum transfers. Substituting h from the budget constraint into the objective function, we can rewrite the seller's CM problem as follows

$$W^s(m) = \max_{x,m_{+1}} U(x) - x + \phi(m - m_{+1}) + \mathcal{T} + \beta V^b_{+1}(m_{+1}) \tag{5}$$

which yields the following first order and envelope conditions

$$U'(x) = 1 \tag{6}$$

$$\beta V^{s'}_{+1}(m_{+1}) = \phi \tag{7}$$

$$W^s_m = \phi. \tag{8}$$

[11] If allowed to participate in the banking sector, sellers will be indifferent in equilibrium as discussed in Rocheteau and Nosal (2017, chap. 8, p. 228). It is enough to assume they face an arbitrarily small cost of accessing banks to rule out their participation.

AM problem: Sellers take the price of the AM good, p, as given and choose the quantity to be supplied, q_s, by solving the following problem

$$V^s(m) = \max_{q_s} \ -c(q_s) + W^s(m + pq_s) \tag{9}$$

which results in the following first order condition

$$c'(q_s) = \phi p. \tag{10}$$

Taking the derivative of the seller's expected AM value function (9) with respect to money holdings, we have that

$$V_m^s = \frac{\partial V^s(m)}{\partial m} = W_m^s = \phi \tag{11}$$

where we replaced W_m^s by its value from (8). This last expression reflects the fact that sellers can only benefit from carrying an additional unit of money by spending it in the next CM. This is the case as they do not consume in AM.

4.2 Buyers' problem

Depending on their idiosyncratic cost of accessing banking services ε, some buyers will find it worthwhile to access the BM to borrow or deposit. In contrast, other buyers will not participate in the BM and will only use their money holdings to consume in the AM. We call the former financially included buyers and the latter financially excluded buyers. In what follows, we characterize the optimal choices of each type of buyer.

4.2.1 Financially excluded buyers

CM problem: The problem facing financially excluded buyers is similar to the one facing buyers in Rocheteau and Wright (2005). At the beginning of the third sub-period, a financially excluded buyer enters a frictionless competitive Walrasian market with m units of fiat money. In this market, buyers choose CM consumption and effort as well as fiat money holdings, m_{+1}, to bring forward to the next period. Formally, buyers solve the following problem

$$W^b(\varepsilon, m) = \max_{x,h,m_{+1}} \ U(x) - h + \beta V(\varepsilon, m_{+1}) \tag{12}$$

$$\text{s.t. } \ x + \phi m_{+1} = h + \phi m + \mathcal{T} \tag{13}$$

where V^b and W^b denote the AM and CM value functions, respectively. Finally, ε represents the buyer's idiosyncratic cost of accessing financial services. For financially excluded buyers, since they choose not to use banking services, this cost is not incurred. For this reason we omit ε from V^b and W^b in what follows.

Substituting h from the budget constraint into the objective function, we can rewrite the agent's second sub-period problem as follows

$$W^b(m) = \max_{x, m_{+1}} \; U(x) - x + \phi(m - m_{+1}) + \mathcal{T} + \beta V^b(m_{+1}) \tag{14}$$

which yields the following first order and envelope conditions

$$U'(x) = 1, \tag{15}$$

$$\beta V^b_{m+1} = \phi, \tag{16}$$

$$W^b_m = \phi. \tag{17}$$

It is worth highlighting that the consumption of the CM good coincides with the efficient allocation.

AM problem: The expected value function of a financially excluded buyer entering the AM with money holdings m is given by

$$V^b(m) = \sigma \left[u(q_b) + W(m - pq_b)\right] + (1 - \sigma) \left[W(m)\right] \tag{18}$$

where q_b is the amount of the AM good demanded by the buyer, that solves the following problem

$$\max_{q_b} u(q_b) + W^b(m - pq_b, 0, 0) \;\; \text{s.t.} \;\; pq_b \leq m \tag{19}$$

which results in the optimality condition

$$\frac{u'(q_b)}{c'(q_s)} = 1 + \frac{\lambda_m}{\phi}. \tag{20}$$

where we replaced ϕp by its value given by equation (10). λ_m is the Lagrange multiplier associated with the cash feasibility constraint, whereby buyers cannot spend more than the amount of cash they have carried into AM. If the cash feasibility constraint does not bind, i.e. $\lambda_m = 0$, then (20) reduces to $u'(q_b) = c'(q_s)$. If $\lambda_m > 0$ then we have that $\frac{u'(q_b)}{c'(q_s)} > 1$ and $q_b = \frac{m}{p}$.

AM envelope condition: We can now solve for the marginal value of bringing an additional unit of money to AM for financially excluded buyers given the terms of trade they face. If we take the derivative of the expected value function of the AM (given by equation (18)) with respect to money holdings, we have that

$$V_m^b = \frac{\partial V^b(m)}{\partial m} = \sigma\left[u'(q_b)\frac{\partial q_b}{\partial m} + W_m^b\left(1 - p\frac{\partial q_b}{\partial m}\right)\right] + (1-\sigma)W_m^b. \tag{21}$$

As long as holding money is costly (i.e. $\gamma > \beta$), we have $q_b = \frac{m}{p}$, hence $\frac{\partial q_b}{\partial m} = \frac{1}{p} = \frac{\phi}{c'(q_s)}$. From (17), we have $W_m^b = \phi$. Using these results, we can simplify (21) to

$$V_m^b = \phi\left[\sigma\frac{u'(q_b)}{c'(q_s)} + (1-\sigma)\right]. \tag{22}$$

4.2.2 Financially included buyers

CM problem: The choices of financially included buyers are similar to Berentsen et al. (2007). They enter CM with a portfolio of fiat money $\hat{m}$, nominal loans ℓ and nominal deposits d. In this market, buyers have to incur a cost ε to gain access to BM in the next period. Given that, they choose their CM consumption and effort as well as their fiat money holdings to bring forward to the next period. Formally, financially included buyers solve

$$\hat{W}^b(\varepsilon, m, \ell, d) = \max_{x,h,\hat{m}_{+1}} \; U(x) - \varepsilon - h + \beta\hat{V}^b(\varepsilon, \hat{m}_{+1}) \tag{23}$$

$$\text{s.t. } \; x + \phi\hat{m}_{+1} = h + \phi\hat{m} + \phi(1+i_d)d - \phi(1+i_\ell)\ell + \mathcal{T} \tag{24}$$

where $\hat{V}^b$ and $\hat{W}^b$ denote the AM and CM value functions of financially included agents, respectively, i_d represents the interest rate earned on deposits and i_ℓ is the lending rate.

Substituting h from the budget constraint into the objective function and solving for the first order and envelope conditions yields

$$U'(x) = 1 \tag{25}$$

$$\beta\hat{V}^b_{\hat{m}_{+1}} = \phi \tag{26}$$

$$\hat{W}^b_{\hat{m}} = \phi \tag{27}$$

$$\hat{W}^b_d = \phi(1+i_d) \tag{28}$$

$$\hat{W}^b_\ell = -\phi(1+i_\ell). \tag{29}$$

AM and BM problem: At the beginning of the period, and before the preference shock is realized, the expected value of a financially included buyer with money holdings $\hat{m}$, is given by

$$\hat{V}^b(\varepsilon, \hat{m}) = \sigma \left[u(\hat{q}_b) + \hat{W}^b(\varepsilon, \hat{m} + \ell - p\hat{q}_b, 0, \ell) \right] + (1-\sigma)\hat{W}^b(\varepsilon, \hat{m} - d, d, 0) \tag{30}$$

where $\hat{q}_b$ is the quantity of AM goods consumed by financially included buyers. After preference shocks are realized, banks open and offer their services. To increase the quantity of AM goods they can purchase, buyers can borrow fiat money to top up their real balances. Alternatively, they can deposit their idle money holdings with the bank and earn some interest. Once banks close their doors, buyers and sellers trade the AM good for fiat money.

A financially included buyer, who does not consume in AM, decides how much he will deposit in the bank by solving

$$\max_{d} \quad \hat{W}^b(\varepsilon, \hat{m} - d, d, 0) \;\text{ s.t. }\; d \leq \hat{m} \tag{31}$$

where the constraint reflects the fact that the buyer cannot deposit more than the fiat money in his possession. Using equations (27) and (28), the first order condition for the choice of deposits is given by $\lambda_d = \phi i_d$ where λ_d is the Lagrange multiplier on the constraint. This implies that this buyer will deposit all his money holdings, i.e. $d = \hat{m}$, as long as money is valued ($\phi > 0$) and the interest rate earned on deposits is positive ($i_d > 0$).

A financially included buyer who wants to consume in AM faces the following problem

$$\max_{\hat{q}_b, \ell} \quad u(\hat{q}_b) + \hat{W}^b(\varepsilon, \hat{m} + \ell - p\hat{q}_b, 0, \ell) \;\text{ s.t. }\; p\hat{q}_b \leq \hat{m} + \ell. \tag{32}$$

Using equations (10), (27) and (29), the optimal choices in AM can be summarized as follows

$$\ell = p\hat{q}_b - \hat{m} \tag{33}$$

$$\frac{u'(\hat{q}_b)}{c'(q_s)} = 1 + \frac{\hat{\lambda}_m}{\phi} \tag{34}$$

$$i_\ell = \frac{\hat{\lambda}_m}{\phi} \tag{35}$$

where $\hat{\lambda}_m$ is the Lagrangian corresponding to the cash feasibility constraint. In this case, it is the amount of cash they have carried from the previous CM and the cash loan they have borrowed from the bank. If the cash feasibility constraint does not bind, such that $\hat{\lambda}_m = 0$, then (34) and (35) reduce to $u'(\hat{q}_b) = c'(q_s)$. If $\hat{\lambda}_m > 0$ then

$$\frac{u'(\hat{q}_b)}{c'(q_s)} = 1 + i_\ell. \tag{36}$$

which states that agents borrow up to the point where the marginal benefit of borrowing is equal to its marginal cost $1 + i_\ell$. Their AM consumption is $\hat{q}_b = \frac{\hat{m}+\ell}{p}$.

AM envelope condition: Having characterized the terms of trade in AM, we can now establish the marginal value of bringing an additional unit of money to AM for financially included buyers. Taking the derivative of $\hat{V}^b$ in equation (30) with respect to money holdings, we have that

$$\begin{aligned}\hat{V}^b_{\hat{m}} = \frac{\partial \hat{V}^b(\varepsilon, \hat{m})}{\partial \hat{m}} &= \sigma \left[u'(\hat{q}_b)\frac{\partial \hat{q}_b}{\partial \hat{m}} + \hat{W}^b_m \left(1 + \frac{\partial \ell}{\partial \hat{m}} - p\frac{\partial \hat{q}_b}{\partial \hat{m}}\right) + \hat{W}^b_\ell \frac{\partial \ell}{\partial \hat{m}} \right] \\ &+ (1-\sigma)\left[\hat{W}^b_m \left(1 - \frac{\partial d}{\partial \hat{m}}\right) + \hat{W}^b_d \frac{\partial d}{\partial \hat{m}} \right].\end{aligned} \tag{37}$$

From (27), (28) and (29) we have $\hat{W}^b_m = \phi$, $\hat{W}^b_d = \phi(1+i_d)$ and $\hat{W}^b_\ell = -\phi(1+i_\ell)$. In addition, we know that $\frac{\partial d}{\partial \hat{m}} = 1$, since the buyer will deposit all his money holdings as long as $i_d > 0$. Taking into account what precedes simplifies the above expression to

$$\hat{V}^b_{\hat{m}} = \sigma \left[u'(\hat{q}_b)\frac{\partial \hat{q}_b}{\partial \hat{m}} + \phi \left(1 + \frac{\partial \ell}{\partial \hat{m}} - p\frac{\partial \hat{q}_b}{\partial \hat{m}}\right) - \phi(1+i_\ell)\frac{\partial \ell}{\partial \hat{m}} \right] + (1-\sigma)\phi(1+i_d). \tag{38}$$

For $i_\ell > 0$, $p\hat{q}_b = \hat{m} + \ell$ which means $p\frac{\partial \hat{q}_b}{\partial \hat{m}} = 1 + \frac{\partial \ell}{\partial \hat{m}}$. Incorporating this and (36) into the previous expression and rearranging terms we get

$$\hat{V}^b_{\hat{m}} = \phi\left[\sigma i_\ell + (1-\sigma)i_d + 1\right]. \tag{39}$$

4.2.3 Decision to access banking services

At the beginning of CM, a buyer facing cost ε decides whether to access financial services in the next period. Formally, this agent solves the following problem

$$\max\{-\phi m_{+1} + \beta V^b_{+1}(m_{+1}), -\varepsilon - \phi \hat{m}_{+1} + \beta \hat{V}^b_{+1}(\varepsilon, \hat{m}_{+1})\} \tag{40}$$

where the first element is independent of ε, while the second is monotonically decreasing in ε. Buyers with cost $\varepsilon \leq \tilde{\varepsilon}$, choose to use banking services, whereas those with $\varepsilon \geq \tilde{\varepsilon}$ choose not to. The cost threshold $\tilde{\varepsilon}$, that determines who will participate in BM, is implicitly defined by the following indifference condition

$$-\phi m_{+1} + \beta V_{m+1}(m_{+1}) = -\tilde{\varepsilon} - \phi \hat{m}_{+1} + \beta \hat{V}^b_{m+1}(\tilde{\varepsilon}, \hat{m}_{+1}). \tag{41}$$

4.3 Banks

A bank accepts nominal deposits, d, paying nominal interest rate i_d, and issues loans ℓ, charging borrowers the nominal interest rate i_ℓ. Banks and buyers trade loans and deposits in perfectly competitive markets. From now on, we restrict our attention to banking systems where a bank can only issue an amount of loans smaller or equal to the amount of deposits it accepts. Since banks face free entry it follows that $i_\ell = i_d \equiv i$.

5 Stationary monetary equilibria

In a stationary monetary equilibrium, we know from equations (7) and (11) that sellers will be indifferent between carrying money across periods or not when $\gamma = \beta$.[12] From now on, we focus only on equilibria where $\gamma > \beta$ such that sellers do not hold any money balances.

Financially excluded buyers' intertemporal equation resulting from (16) and (22) is given by

$$\bar{\iota} = \sigma \left[\frac{u'(q_b)}{c'(q_s)} - 1 \right] \tag{42}$$

where $\bar{\iota} \equiv \frac{\gamma - \beta}{\beta}$ represents the opportunity cost of holding money from one CM to the next. The right hand side of equation (42) represents the expected net return: an extra unit of money allows a financially excluded buyer to consume an extra unit of the AM good with probability σ.

The intertemporal trade-off facing financially included buyers results from equations (26) and (39) and can be summarized as

$$\bar{\iota} = \sigma i_\ell + (1 - \sigma) i_d \tag{43}$$

where

$$i_\ell = \frac{u'(\hat{q}_b)}{c'(q_s)} - 1 \tag{44}$$

holds from equation (36). The left hand side of equation (43) describes the net cost of holding one extra unit of money to the next period while the right hand side represents the net expected return. An extra unit of money allows a borrower to reduce his costs by borrowing one unit less of money from the banking sector at cost i_ℓ. For the depositor, taking one extra unit of money allows him to deposit it and earn return i_d.

[12] More precisely, if $\gamma = \beta$ holds, sellers will carry an indeterminate amount of money as discussed by Rocheteau and Wright (2005).

Using the free entry condition in the banking sector, $i_\ell = i_d \equiv i$, in (43) and (44) we get

$$\bar{\iota} = i \tag{45}$$

and

$$i = \frac{u'(\hat{q}_b)}{c'(q_s)} - 1 \tag{46}$$

where i, the nominal interest rate prevailing in the BM, equals in equilibrium the opportunity cost of real balances $\bar{\iota}$. Comparing equations (42) and (46) indicates that the quantity consumed by financially included buyers is always higher than the quantity consumed by financially excluded buyers since the former face a lower marginal cost of carrying money balances. This is due to the interest rate they can earn on their idle money balances.

To close the model, markets have to clear. In particular, the amount of goods traded in the AM has to satisfy

$$\sigma\left((1 - F(\tilde{\varepsilon}))q_b + F(\tilde{\varepsilon})\hat{q}_b\right) = q_s, \tag{47}$$

while BM clearing requires

$$\sigma F(\tilde{\varepsilon})\ell = (1 - \sigma)F(\tilde{\varepsilon})d. \tag{48}$$

Using the equilibrium conditions above, we can simplify the cost threshold equation (A.5), derived in the appendix, to get

$$\tilde{\varepsilon} = \beta\sigma\left[(u(\hat{q}_b) - u'(\hat{q}_b)\hat{q}_b) - (u(q_b) - u'(q_b)q_b)\right] \tag{49}$$

which states that the degree of financial inclusion is determined by the discounted net utility gain of accessing banking services weighted by the probability of the preference shock.

Finally, the money demanded by buyers equals the money supplied by the government such that

$$M = (1 - F(\tilde{\varepsilon}))m + F(\tilde{\varepsilon})\hat{m}. \tag{50}$$

Definition 1 *Given $\bar{\iota}$, a monetary equilibrium consists of a threshold $\tilde{\varepsilon}$, an interest rate on loans and deposits i, money holdings $\{m, \hat{m}\}$, AM quantities $\{q_b, \hat{q}_b, q_s\}$ and prices $\{\phi, p\}$ that together satisfy the optimal choices of agents and clear markets.*

The equations for a monetary equilibrium with competitive banks are listed in Appendix A.2. In what follows we discuss the monetary equilibria resulting from different values taken by the model's parameters and in particular the money growth rate γ and the distribution of financial access costs $F(\varepsilon)$.[13]

[13] We always operate under the assumption $\sigma < 1$.

5.1 Pure monetary equilibria

When money is costless to hold and/or when accessing banking services is too costly, the environment exhibits monetary equilibria where there is no demand for financial services.

Proposition 1 *At the Friedman rule $\gamma = \beta$, when the costs of accessing financial services are strictly positive buyers choose to remain financially excluded. When they are costless, buyers are indifferent. Thus, banking services are irrelevant for the real allocations and the equilibrium consumption coincides with the first-best allocation:* $q_b = \hat{q}_b = q_b^*$.

Proof. A proof can be found in Appendix B.1. ■

When the Friedman rule is satisfied ($\gamma = \beta$), carrying money across periods is costless. This means the risk that real balances remain unused following a negative preference shock is irrelevant. Buyers can perfectly self-insure and there is no demand for financial intermediation.

Proposition 2 *When $\gamma > \beta$ and the costs of financial access are sufficiently high, buyers choose to remain financially excluded* ($F(\tilde{\varepsilon}) = 0$). *In this case, there exists a unique monetary equilibrium where BM is inactive and consumption is below the first-best:* $q_b < q_b^*$.

Proof. A proof can be found in Appendix B.2. ■

It is worth highlighting that the equilibrium described in Proposition 1 corresponds to the monetary equilibrium in Rocheteau and Wright (2005) under the Friedman rule, while Proposition 2 corresponds to the monetary equilibrium away from the Friedman rule.

5.2 Monetary and banking equilibria

Here we analyze situations where equilibria with both financial intermediation and money can coexist.

Proposition 3 *When $\gamma > \beta$ and the costs of financial access are not too high, a monetary equilibrium with limited BM participation* ($0 < F(\tilde{\varepsilon}) < 1$) *exists.*

Proof. A proof can be found in Appendix B.3. ■

The previous proposition states that when carrying real balances across periods is costly and the cost of accessing banking services are not too high, a stationary monetary equilibrium exists where a measure of buyers choose to access the BM, while the rest choose not to.

Proposition 4 *If $\gamma > \beta$ and $\underline{\varepsilon} = \bar{\varepsilon} = 0$, a unique monetary equilibrium with full BM participation exists. The equilibrium consumption is below the first-best allocation i.e. $\hat{q}_b < q_b^*$.*

Proof. A proof can be found in Appendix B.4. ■

When it is costless to access financial services, all buyers choose to be financially included and the resulting equilibrium is the one in Berentsen et al. (2007).

6 Equilibrium properties

In this section, we focus on monetary equilibria where participation in the BM is limited, which corresponds to the equilibria described in Proposition 3. This type of equilibria is of particular interest for two reasons. First, they describe what we observe in developing economies. Second, from a theoretical point of view, these equilibria involve two types of agents facing different liquidity constraints. This situation results in interesting interactions and non-trivial inefficiencies that can help explain some of the consumption inequality observed in developing countries.

Before we delve into the results, it is worth mentioning that any parameter or policy change that affects the trade-off facing buyers, when deciding whether to access banks or not, can have important consequences for welfare. Notice that when a buyer chooses to access banking services, he does not internalize the impact of his decision on the price of the AM goods. This pecuniary externality (a situation in which the action of an agent affects another agent only through its effect on prices) always occurs in models where agents trade in a competitive market or more generally when prices faced by an agent depend on the choices of other agents.[14] As emphasized by Loong and Zeckhauser (1982) and Greenwald and Stiglitz (1986), in an economy with complete markets, pecuniary externalities do not generate inefficiencies.[15] However, when agents face incomplete markets, pecuniary externalities can result in substantial inefficiencies.[16] In our setting, all agents face liquidity constraints. In addition, some are excluded from participation in the credit market, which makes liquidity constraints differ among agents. An increase in the AM good's price resulting from higher financial inclusion can tighten buyers' liquidity constraint. This can have potentially a stronger effect on the financially excluded. In this case, the welfare losses of these agents

[14] Pecuniary externalities do not occur in markets where prices faced by individual agents are independent of aggregate quantities, e.g. a market with bilateral trades where prices are bargained over between the two parties.

[15] As opposed to technological externalities which usually result in inefficiencies.

[16] For a related discussion in models with financial frictions see Dávila and Korinek (2017). Moreover, generically the direction of the inefficiency cannot be predicted (Loong and Zeckhauser, 1982; Dávila and Korinek, 2017).

Table 1: Parameter values

Parameter	Description	Value
β	Discount factor	0.95
γ	Growth rate of money	1.03
A	Parameter of AM utility function	1.20
a	Parameter of AM utility function	0.20
α	Parameter of AM cost function	0.10
σ	Probability of consuming in AM	0.50
B	Probability of CM utility	1.00
$F(\varepsilon)$	Distribution of financial access costs	Uniform [0,1]

might not be compensated by the welfare gains of the others and hence aggregate welfare might decrease.

To shed more light on the equilibrium properties of this model, in what follows we consider specific functional forms and parameter values. In particular, the AM utility and cost functions are given by $u(q) = A\frac{q^{1-a}}{1-a}$ and $c(q) = \frac{q^{1+\alpha}}{1+\alpha}$, respectively. The CM utility is given by $U(x) = B \log x$. We solve the model numerically using standard parameter values from the literature, which are summarized in Table 1. In a later section, we conduct a thorough calibration exercise for the case of India to better discipline our choice of parameters.

Result 1 *Consumption of financially excluded buyers in AM, q_b, is decreasing in the money growth rate γ and increasing in the preference shock σ. The AM consumption of financially included buyers $\hat{q}_b$ is ambiguous in γ and decreasing in σ.*

The left-hand side panel of Figure 1 illustrates Result 1. A higher money growth rate γ increases the marginal cost of holding money across periods which makes AM consumption more costly and reduces real balances. This lowers the quantity consumed q_b and $\hat{q}_b$ as well as aggregate supply q_s. Maintaining the assumption of a strictly convex cost function in AM ($c''(q_s) > 0$), lower q_s reduces the AM price. This partially compensates buyers against the higher inflation. Since financially included buyers face a lower marginal cost of consumption in the AM, an increase in γ affects them less relative to financially excluded buyers. For the former, the fall in the AM price can be so strong that it dominates the cost of higher inflation resulting in an increase in $\hat{q}_b$. This is the case when the share of financially excluded agents and the liquidity risk are very high. However, aggregate welfare is invariantly decreasing in γ.

Result 1 also highlights the effect of changes in the liquidity risk resulting from the preference shock σ. An increase in σ implies a higher probability of AM consumption and hence a higher return of holding money across periods. As a consequence q_b is higher. In

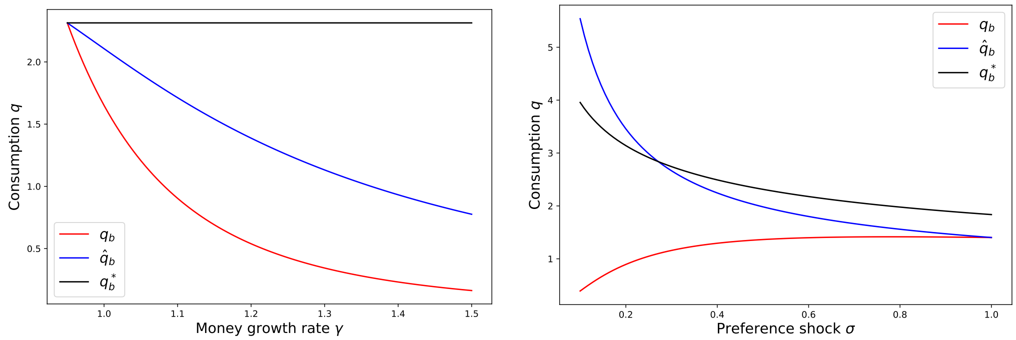

Figure 1: AM consumption as a function of γ and σ

contrast, financially included buyers are perfectly compensated against the liquidity risk through banking services. This means that changes in σ do not affect $\hat{q}_b$ directly. Nevertheless, a general equilibrium effect takes place whereby higher q_b increases the price of the AM good and reduces $\hat{q}_b$. A second general equilibrium effect works through the extensive margin. An increase in σ reduces $F(\tilde{\varepsilon})$ and hence puts downward pressure on the AM good price. This in turn increases both q_b and $\hat{q}_b$. However, the former effect always dominates and $\hat{q}$ is always decreasing in σ.

Result 2 *A high liquidity risk (low σ) may result in over-consumption by the financially included buyers such that $\hat{q}_b > q_b^*$.*

Result 2 is illustrated in the right hand side panel of Figure 1. Result 2 shows that for very low values of σ, the liquidity constraint on financially excluded buyers can be so tight that it greatly lowers q_b and q_s. As a result, the AM price sufficiently adjusts to push $\hat{q}_b$ above the socially efficient quantity q_b^*.

Result 3 *Consumption inequality measured by the ratio $\dfrac{\hat{q}_b}{q_b}$ is increasing in inflation γ and decreasing in the preference shock σ.*

Proof. The proof is available in Appendix B.5. ■

Figure 2 is a direct illustration of Result 3. It depicts the ratio of consumption $\frac{\hat{q}_b}{q_b}$ as a function of the money growth rate γ and the preference shock σ. Intuitively, buyers who access banking services are compensated against the risk of idle money balances through the interest they earn on deposits as opposed to financially excluded buyers. This limited access to banking results in different returns on holding money across periods. An increase in the money growth rate, γ, will then have a stronger effect on q_b when compared to $\hat{q}_b$. As a result, we observe an increase in the ratio of AM consumption of financially included to excluded buyers.

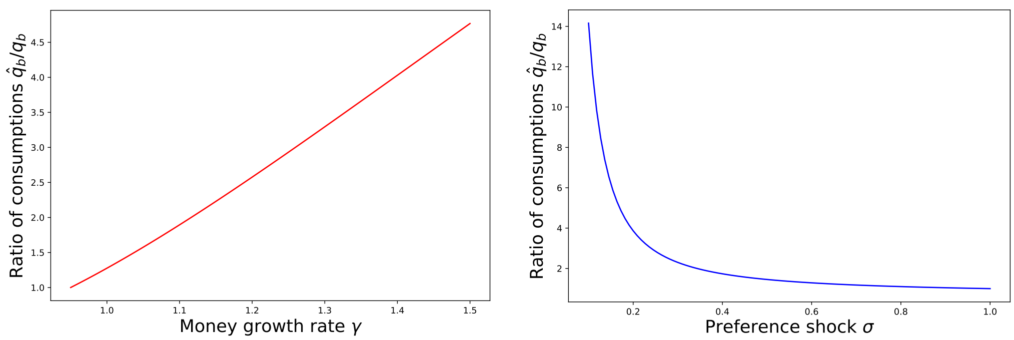

Figure 2: Ratio of AM consumptions as a function of γ and σ

A lower liquidity risk arising from the preference shock σ directly reduces consumption inequality as depicted in the right panel of Figure 2. This is because access to banking services is the only source of inequality in our model. Reducing the liquidity risk makes access to banks less relevant for AM consumption and hence reduces consumption inequality.

Finally, under the parametrization given in Table 1, we are able to establish the following result.

Result 4 *The measure of financially included buyers $F(\tilde{\varepsilon})$ is non-monotonic in the rate of money growth γ and the preference shock σ. Moreover, for a given $\sigma \in (0, 1)$, there exists $\gamma > \beta$ such that financial inclusion is maximized.*

Figure 3 illustrates Result 4. The left panel depicts the measure of financially included buyers $F(\tilde{\varepsilon})$ as a function of the money growth rate γ. Changes in $F(\tilde{\varepsilon})$ result from the differential effect of changes in inflation on the two types of buyers as discussed above. As the economy moves away from the Friedman rule, financial inclusion increases, reaches a maximum and then decreases. To understand this non-monotone relationship, it is worth emphasizing that the interest rate earned on deposits compensates depositors on their idle money balances. Given that, an increase in inflation has two effects. First, it reduces the demand for real money balances and hence AM consumption. Second, it increases the value of accessing banks to earn interests on idle money balances. For low levels of inflation and nominal rates, the opportunity cost of holding money is small. This makes the risk of ending with idle money balances negligible for buyers. As a consequence, only those with a low ε are willing to pay to compensate against this risk. As inflation increases, the opportunity cost of holding money increases, which in turn makes the risk of ending with idle money balances more costly. As a result, more buyers are willing to pay the cost of financial access. At the same time, higher inflation reduces the holdings of real money balances. When inflation is sufficiently high, the quantity of real balances held is so small that the gain from accessing

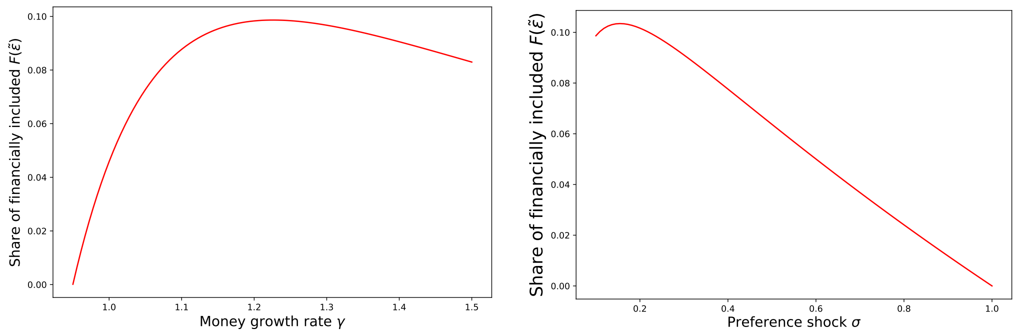

Figure 3: Financial inclusion as a function of γ and σ

banking services does not justify the payment of the cost ε.[17] This means that buyers with high values of ε start dropping out from the banking sector.

This extensive margin effect of inflation is not present in models where the measure of agents with access to financial markets is exogenous. Changes in $F(\tilde{\varepsilon})$ produce a pecuniary externality, which can have an additional effect on welfare. In the region of the parameter space where an increase in γ results in higher $F(\tilde{\varepsilon})$, the higher demand from the new financially included buyers puts upward pressure on the price of the AM good, which further reduces q_b and $\hat{q}_b$. In the region of the parameter space where an increase in γ results in a decrease in $F(\varepsilon)$, the pecuniary externality operates in the opposite direction.

The impact of an increase in the liquidity risk is shown in the right panel of Figure 3. The relationship is also non-monotone as the previous logic applies. When the probability of a negative preference shock increases, i.e. lower σ, the share of financially included buyers increases. However, as the liquidity risk becomes very high, i.e. very low σ, buyers with high banking costs, don't find it worthwhile to pay ε to access banks as the amount of real balances they hold is small.

7 Quantitative analysis

7.1 Calibration

In order to evaluate the impact of various policies on financial inclusion, welfare and inequality as well as provide more discipline when choosing the parameter values, we calibrate the model to the Indian economy. We consider annual data over the period 1970 to 2016.[18]

We assume the utility of consumption takes the form $u(q) = A\frac{q^{1-a}}{1-a}$ for the AM good and $U(x) = B\log x$ for the CM good. The AM cost function is set as $c(q) = \frac{q^{1+\alpha}}{1+\alpha}$. We

[17] The cost of financial access ε is incurred in real terms.

[18] All the data used is publicly available from the Reserve Bank of India.

Table 2: Calibration results

Parameters			
σ	Preference shock		0.643
a	Parameter of AM utility		0.013
A	Parameter of AM utility		2.528
α	Parameter of AM production function		2.649
B	Parameter of CM utility		7.629
Targets		**Data**	**Model**
$L(\bar{i})$	Average Money Demand	0.175	0.175
ϵ	Elasticity of Money Demand to i	-0.195	-0.195
$F(\tilde{\varepsilon})$	Share of financially included buyers	0.352	0.352
Validation		**Data**	**Model**
Consumption Gini coefficient		0.38	0.11
Demand deposits to M1 ratio		45.3%	32.5%
Non-durables related consumer credit to GDP		13.0%	6.3%

divide the model's parameters into two groups: independent parameters, $\{\beta, \gamma, F, \underline{\varepsilon}, \bar{\varepsilon}\}$, and jointly calibrated parameters, $\{\sigma, A, a, \alpha, B\}$. We set the money growth rate γ to match the average annual change in the Indian CPI over the period 1970 to 2016 of 7.92%. For the nominal interest rate, i, we consider average call money interest rate, which is 8.67%. This then implies an average annual real interest rate of 0.75% and a discount factor β of 0.993.

Since direct measurements for the distribution of the cost of accessing banking services, $F(\varepsilon)$, are not available, we use the per capita density of bank branches across Indian states as a proxy.[19] This proxy implicitly assumes that a higher per capita number of bank branches implies a lower cost of accessing financial services, everything else being equal.[20] Given our proxy for financial inclusion, we define ε as the inverse of bank branch density and use maximum likelihood to fit the parameters of a log-normal distribution. This results in a location parameter of 0, a scale of 0.129 and a standard deviation of 0.341.

The second group of parameters is calibrated jointly by matching three empirical moments: the average money demand, the interest rate elasticity of money demand and the degree of financial inclusion. The first two targets are used to fit the empirical money demand.[21]

[19] Detailed data on bank branches per state and district is publicly available from the Reserve Bank of India.

[20] We refer the reader to our working paper version where we provide cross-country evidence that financial inclusion is positively correlated with bank branch density.

[21] This is the procedure used in Lagos and Wright (2005). Alternatively, one can use all the information in the data by solving for the parameter values which minimize the distance between the model-based money demand and the observed money demand for each observed interest rate. However this procedure is computationally more expensive and doesn't change the calibration results significantly.

In our model, the money demand relationship is given by

$$L \equiv \frac{M}{PY} = \frac{F(\tilde{\varepsilon})\phi\hat{m} + (1 - F(\tilde{\varepsilon}))\phi m}{\phi p q_s + 2B} \tag{51}$$

where the numerator represents the sum of real balances demanded by both financially included and excluded buyers, while the denominator represents the sum of output produced in the AM and CM. To obtain the empirical estimates of the two money demand related moments we run the following regression

$$\log M_t/P_tY_t = \beta_1 + \beta_2 \log i_t + \nu_t. \tag{52}$$

using annual data over the period 1970 to 2016. The empirical counterpart of M is M1, which includes currency and demand deposits but excludes time deposits. For the rest of the observables, we use nominal GDP as a measure of PY and the call money rate for i.[22] The OLS estimate for β_2 is used as a target for the model-based interest rate elasticity of money demand, $\frac{\partial L(i)/L(i)}{\partial i/i}$, while the average money demand in the data is used as a target for the level of money demand L at the average i.

To match the observed degree of financial inclusion, we target the percentage of the Indian adult population reported to own an account at a formal financial institution, which was 35.2% in 2011.[23] The results of the joint calibration are presented in Table 2. As we can see, the model is able to exactly match the targeted moments using the calibrated parameters.

Using the benchmark calibration, we now examine the performance of our model by comparing moments that have not been targeted. In particular, we consider the Gini coefficient, the demand deposits to M1 ratio and the share of intra-period credit to total consumption.[24] The comparisons with the data counterparts are reported in the lower part of Table 2. Our simple model is able to explain around a third of the observed consumption inequality in India. This is surprising since inequality, in our model, is only due to differences in the cost of accessing banking services. The model is also able to explain two thirds of the average ratio of demand deposits to M1. If we interpret the AM as consumption of non-durables, we can also explain around half of the average share of non-durable goods in consumer credit to GDP in India. We interpret these results as showing that the channel explored in this paper, the choice of participating in the banking system, is a relevant one.

[22] Due to limited data availability, we use the call money rate instead of the commercial paper rate used in the literature.

[23] This statistic is taken from the World Bank's Global Findex database.

[24] As a measure of consumption inequality, we use the consumption Gini coefficient (cf. Appendix A.3 for details). The Gini coefficient takes into account both the intensive and extensive margins.

Table 3: Effect of reducing steady state inflation $\pi = 7.9\%$

	$\Delta F(\tilde{\varepsilon})$	Δq_b	$\Delta \hat{q}_b$	Δq_s	$\Delta \mathcal{W}$	Δ Gini
$\pi = 6.9\%$	-1.11 pp	42.97%	0.59%	0.35%	0.02%	-0.13 pp
$\pi = 2.9\%$	-16.84 pp	560.14%	4.39%	1.79%	0.15%	-2.49 pp

7.2 Numerical experiments

7.2.1 Inflation

Table 3 presents the results of a 1 percentage point (pp) and 5 pp reduction in the rate of inflation $\pi \equiv 1 - \gamma$ relative to the benchmark calibration. Lowering inflation reduces the measure of financially included buyers. This implies that the benchmark steady state is on the upward sloping part of the relationship between financial inclusion and inflation seen in the left panel of Figure 3. Lowering inflation increases the consumption of both types. However, financially excluded buyers enjoy a larger proportional increase as they are more liquidity constrained. As a result, we observe a reduction in inequality and an increase in welfare.[25] In particular, decreasing inflation by 5 pp increases welfare by 0.15% and reduces consumption inequality by 2.49 pp.

7.2.2 Costs of financial access

India implemented an aggressive policy of mandating banks to increase the number of their branches in remote areas starting from 2010. This policy aimed at lowering the cost of financial access. Not surprisingly, the 2014 World Bank's survey reports that India had a degree of financial inclusion of 52.8% compared to 35.2% in 2011.[26]

To capture the impact of this policy, we re-estimate the distribution $F(\varepsilon)$ using 2014 data on bank branches and the state-level population growth rates from 2011 to 2014. We then use this new distribution to calculate a new steady state equilibrium, while keeping the rest of the parameters at their benchmark values. We find that our model delivers a degree of financial inclusion equal to 42%, capturing 38.9% of the observed change. Other equilibrium quantities are presented in Table 4. In particular, welfare increases by 0.03%, while consumption inequality falls by 0.52 pp.

[25] Our welfare measure is given by $\mathcal{W} = \sigma F(\tilde{\varepsilon})u(\hat{q}_b) + \sigma(1 - F(\tilde{\varepsilon}))u(q_b) - c(q_s) + 2U(x^*) - 2x^* - \int_{\underline{\varepsilon}}^{\tilde{\varepsilon}} \varepsilon \, dF(\varepsilon)$ where the integral represents the sum of financial access costs incurred by financially included buyers.

[26] These numbers exclude access to mobile banking which was reported by only 2.4% of the adult population in 2014. Statistics on mobile banking were not collected in the 2011 survey.

Table 4: Effects of changing $F(\varepsilon)$

	$\Delta F(\tilde{\varepsilon})$	Δq_b	$\Delta \hat{q}_b$	Δq_s	$\Delta \mathcal{W}$	Δ Gini
F_{2014}	6.8 pp	-14.93%	-14.93%	0.08%	0.03%	-0.52 pp

7.3 Financial inclusion policies

In this section we discuss various policy instruments aimed at increasing financial inclusion and analyze their impact on welfare and consumption inequality.

7.3.1 Transfer to bank account holders

We first consider direct monetary transfers to the owners of bank accounts. In practice, this policy can be implemented either by changing pre-existing cash-based government transfers into bank account transfers or through a new subsidy scheme. In the former case, it allows the government to leverage the disbursement of pensions and social benefits in order to increase financial inclusion. Paying these transfers through bank accounts might represent the first contact with the formal financial sector for a large part of the population. Similar policies have been implemented in Brazil, Colombia, India, Mexico and South Africa (Pickens et al., 2009). Worth mentioning is India's Direct Benefit Transfer (DBT) program, operating since 2013. It is recognized as the world's largest targeted benefit transfer scheme. The DBT program pays government subsidies directly to the beneficiaries through their bank accounts with the objective of increasing financial inclusion. This has the additional advantage of reducing corruption by bypassing the handling of cash by government officials.

Let us define the monetary subsidy by τ. We assume that such subsidy is the same for both depositors and borrowers. Monetary transfers are first given to banks, which subsequently pass them to all agents who use their financial services. We assume this subsidy is financed through lump-sum taxes in CM. Note then that τ enters directly the AM value function of financially included buyers (30), which is now given by

$$\hat{V}^b(\varepsilon, \hat{m}) = \sigma \left[u(\hat{q}_b) + \hat{W}^b(\varepsilon, \hat{m} + \tau + \ell - p\hat{q}_b, 0, \ell) \right] + (1-\sigma)\hat{W}^b(\varepsilon, \hat{m} + \tau - d, d, 0). \quad (53)$$

The amount deposited by liquidity-unconstrained buyers is now $d = \hat{m} + \tau$, while the amount of loans that the liquidity-constrained buyers desire is $\ell = p\hat{q} - \hat{m} - \tau$. Note that this transfer scheme affects directly the cost threshold, $\tilde{\varepsilon}$, which is now given by

$$\tilde{\varepsilon} = \gamma\phi\tau + \beta\sigma \left[(u(\hat{q}_b) - u'(\hat{q}_b)\hat{q}_b) - (u(q_b) - u'(q_b)q_b) \right]. \quad (54)$$

As one would expect, the equilibrium money holdings of financially excluded buyers remain unchanged, while the financially included buyers reduce their balances by exactly the amount of the transfer. Thus, we have that $\hat{m} = \sigma p \hat{q}_b - \tau$. The rest of the equilibrium conditions are the same as before.

Changes in τ translate into a pure shift in the share of financially included buyers $F(\tilde{\varepsilon})$. This shift affects welfare both directly and through the pecuniary externality previously discussed. We know that an increase in $F(\tilde{\varepsilon})$ means that some buyers shift their consumption from q_b to a higher quantity $\hat{q}_b$ with a positive effect on aggregate welfare. However, for strictly convex AM cost functions, this higher demand for AM goods puts upward pressure on prices, which results in a decrease in both q_b and $\hat{q}_b$ and has thus a negative effect on welfare. The net welfare effect of the subsidy will depend on which of the two channels dominates. When the pecuniary externality is not present (e.g. with a constant marginal cost of production in the AM) welfare invariably increases.

7.3.2 Interest rate subsidies

We consider two additional policy instruments: a proportional subsidy τ^d to the interest rate earned by depositors and a proportional subsidy τ^ℓ to the interest rate paid by borrowers. These subsidies are financed through lump-sum taxes in the CM. They affect both the intertemporal equilibrium condition of financially included agents, which is now given by

$$\frac{u'(\hat{q}_b)}{c'(q_s)} - 1 = i(1 - \tau^\ell), \tag{55}$$

and the interest rate on deposits and loans is now

$$i(\sigma(1 - \tau^\ell) + (1 - \sigma)(1 + \tau^d)) = \bar{\iota}, \tag{56}$$

while the rest of the equilibrium conditions are the same as in the baseline model. Notice that interest rate subsidies do not directly affect financially excluded buyers. The only effect comes from the pecuniary externality, i.e. changes in the price of the AM good as a response to changes in the quantities demanded by financially included buyers. In contrast, financially included buyers are affected both directly and indirectly by the two instruments. The borrowing rate subsidy τ^ℓ affects the quantity consumed by borrowers directly by reducing the wedge $1 + i(1 - \tau^\ell)$ between their marginal cost and their marginal utility of consumption. There is also an indirect one, through general equilibrium effects, that increases the demand for loans and results in a higher i. The deposit rate subsidy τ^d affects consumption only through general equilibrium effects by increasing the supply of deposits which decreases i.

Table 5: Effect of increasing financial inclusion by 1 pp using different policies

	Δ	Δq_b	$\Delta \hat{q}_b$	Δq_s	$\Delta \mathcal{W}$	Δ Gini
$\phi\tau$	0.004	-2.50%	-2.50%	0.01%	0.004%	-0.07 pp
τ_d	0.199	-33.84%	-0.35%	0.20%	-0.009%	0.13 pp
τ_ℓ	0.166	-33.84%	-0.35%	0.20%	-0.009%	0.13 pp
γ	0.016	-41.55%	-0.64%	-0.54%	-0.031%	0.10 pp

7.3.3 Policy evaluation

Using the benchmark calibrated parameter values, we calculate the welfare and inequality consequences of the previous policies. Table 5 presents the implications for consumption, output, welfare and consumption inequality of increasing financial inclusion by 1 pp under the various policies. As expected, an increase in financial inclusion, i.e. change in the extensive margins, affects the intensive margins by reducing consumption of both included and excluded buyers as explained above. In the case of bank transfers, the decrease in consumption is the same for both financially included and excluded buyers. This is because this policy affects only the credit market participation margin without generating other distortions. In contrast, policies based on interest rate subsidies result in a much higher decrease in consumption for financially excluded buyers, which tends to increase inequality. This is the case as interest rate subsidies not only increase participation in the credit market but also encourage buyers who already had access to banks to consume more which exacerbates the effect of the pecuniary externality on the financially excluded. If the objective of policymakers is to just increase participation in the banking sector, without distorting the incentives of banked individuals, the direct benefit transfer is the most desirable policy.

Using inflation to increase financial inclusion by 1 pp requires the central bank to increase the money growth rate by 1.6 pp. Such policy results in lower consumption for both financially included and excluded agents, lower welfare and higher inequality. Since it directly reduces the consumption of both groups, as opposed to interest rate subsidies, inflation results in a lower increase in inequality. However, the reduction in welfare is almost three times larger.

8 Conclusion

In this paper we present a theoretical framework that sheds some light on financial inclusion and how it relates to welfare and consumption inequality. We consider an environment where informational frictions make money essential. Agents face idiosyncratic preference shocks, inducing heterogeneous liquidity needs. Banks operate as intermediaries in the credit market by accepting deposits and offering loans. To access financial services, agents need to pay a fixed cost which differs across the population. Agents who choose to access banks are compensated against the risk of idle money balances by earning interests on their deposits. In contrast, agents who choose to be financially excluded are not. Hence, the latter get a lower return from holding money across periods. The model exhibits significant consumption inequality as a result of the difference in the access to basic banking services. Moreover, equilibria with over-borrowing and over-consumption, above the socially efficient level, can arise when financially excluded agents are severely liquidity-constrained. We study various policies to increase financial inclusion. Depending on the policy used, it turns out that higher access to banking services can lead to a decrease in welfare and an increase in consumption inequality. This is a direct result of the pecuniariy externality caused by agents' participation decision. An increase in the proportion of agents with access to the banks increases the quantities demanded and hence the price of goods. This price increase tightens the liquidity constraint on all agents and reduces their consumption. This is particularly problematic for agents without access to banks who see their consumption decrease the most. In this case, the aggregate welfare gain from a higher access to banking services might not be enough to compensate against the loss from lower individual consumption.

Our paper shows how financial inclusion is a complex phenomenon that depends, among other factors, on the cost of financial access, on monetary policy and the degree of liquidity risks. Moreover it presents several testable implications that we think are worth examining. As far as monetary policy is concerned, the level of financial inclusion can have non-trivial consequences for central banks' stabilization policies. We leave this issue and others to future research.

References

Albanesi, S. (2007). Inflation and inequality. *Journal of Monetary Economics 54*(4), 1088–1114.

Allen, F., A. Demirgüç-Kunt, L. Klapper, and M. S. M. Pería (2016). The foundations of financial inclusion: Understanding ownership and use of formal accounts. *Journal of Financial Intermediation*.

Allen, F. and D. Gale (1994). Limited market participation and volatility of asset prices. *American Economic Review 84*(4), 933–55.

Beck, T., A. Demirguc-Kunt, and R. Levine (2005). Smes, growth, and poverty: cross-country evidence. *Journal of economic growth 10*(3), 199–229.

Beck, T., A. Demirguc-Kunt, and M. S. M. Peria (2007). Reaching out: Access to and use of banking services across countries. *Journal of Financial Economics 85*(1), 234–266.

Beetsma, R. M. and F. Van Der Ploeg (1996). Does inequality cause inflation?: The political economy of inflation, taxation and government debt. *Public Choice 87*(1-2), 143–162.

Bencivenga, V. R. and G. Camera (2011). Banking in a matching model of money and capital. *Journal of Money, Credit and Banking 43*, 449–476.

Berentsen, A., G. Camera, and C. Waller (2007). Money, credit and banking. *Journal of Economic Theory 135*(1), 171–195.

Berentsen, A., S. Huber, and A. Marchesiani (2014). Degreasing the wheels of finance. *International economic review 55*(3), 735–763.

Burgess, R. and R. Pande (2005, June). Do rural banks matter? evidence from the indian social banking experiment. *American Economic Review 95*(3), 780–795.

Chatterjee, S. and D. Corbae (1992). Endogenous market participation and the general equilibrium value of money. *Journal of Political Economy 100*(3), 615–646.

Chiu, J., M. Dong, and E. Shao (2018). On the welfare effects of credit arrangements. *International Economic Review*.

Chiu, J. and C. A. Meh (2011). Financial intermediation, liquidity, and inflation. *Macroeconomic Dynamics 15*(S1), 83–118.

Cysne, R. P., W. L. Maldonado, and P. K. Monteiro (2005). Inflation and income inequality: A shopping-time approach. *Journal of Development Economics 78*(2), 516–528.

Dávila, E. and A. Korinek (2017). Pecuniary externalities in economies with financial frictions. *The Review of Economic Studies*, rdx010.

Demirgüç-Kunt, A., L. F. Klapper, D. Singer, and P. Van Oudheusden (2015). The global findex database 2014: Measuring financial inclusion around the world.

Diamond, D. W. and P. H. Dybvig (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy 91*(3), 401–419.

Easterly, W. and S. Fischer (2001). Inflation and the poor. *Journal of Money, Credit and Banking*, 160–178.

Erosa, A. and G. Ventura (2002). On inflation as a regressive consumption tax. *Journal of Monetary Economics 49*(4), 761–795.

Gomis-Porqueras, P. (2001, July). When Should Bank Regulation Favor the Wealthy? *Journal of Economic Theory 99*(1-2), 327–337.

Greenwald, B. C. and J. E. Stiglitz (1986). Externalities in economies with imperfect information and incomplete markets. *The Quarterly Journal of Economics 101*(2), 229–264.

Gu, C., F. Mattesini, and R. Wright (2016). Money and credit redux. *Econometrica 84*(1), 1–32.

Lagos, R., G. Rocheteau, and R. Wright (2017). Liquidity: A new monetarist perspective. *Journal of Economic Literature 55*(2), 371–440.

Lagos, R. and R. Wright (2005). A unified framework for monetary theory and policy analysis. *Journal of Political Economy 113*(3), 463–484.

Laha, A., D. Kuri, and P. Kumar (2011). Determinants of financial inclusion: A study of some selected districts of west bengal, india. *Indian journal of finance 5*(8), 29–36.

Levine, R. (2005). Finance and growth: theory and evidence. *Handbook of Economic Growth 1*, 865–934.

Loong, L. H. and R. Zeckhauser (1982). Pecuniary externalities do matter when contingent claims markets are incomplete. *The Quarterly Journal of Economics 97*(1), 171–179.

Lotz, S. and C. Zhang (2016). Money and credit as means of payment: A new monetarist approach. *Journal of Economic Theory 164*, 68–100.

Marshall, J. N. (2004). Financial institutions in disadvantaged areas: a comparative analysis of policies encouraging financial inclusion in britain and the united states. *Environment and Planning A 36*(2), 241–261.

Mehrotra, A. and J. Yetman (2015). Financial inclusion-issues for central banks.

Mehrotra, A. N. and J. Yetman (2014). Financial inclusion and optimal monetary policy.

Menna, L. and P. Tirelli (2017). Optimal inflation to reduce inequality. *Review of Economic Dynamics 24*, 79–94.

Monnet, C. and W. Roberds (2008). Optimal pricing of payment services. *Journal of Monetary Economics 55*(8), 1428–1440.

Pickens, M., D. Porteous, and S. Rotman (2009). Banking the poor via g2p payments. *CGAP Focus Note, World Bank* (58).

Rocheteau, G. and E. Nosal (2017). *Money, Payments, and Liquidity*. MIT Press.

Rocheteau, G. and R. Wright (2005). Money in search equilibrium, in competitive equilibrium, and in competitive search equilibrium. *Econometrica 73*(1), 175–202.

Rojas Breu, M. (2013). The welfare effect of access to credit. *Economic Inquiry 51*(1), 235–247.

Romer, C. D. and D. H. Romer (1999). Monetary policy and the well-being of the poor. *Economic Review, Federal Reserve Bank of Kansas City 84*(1), 21.

Sanches, D. (2011). A dynamic model of unsecured credit. *Journal of Economic Theory 146*(5), 1941–1964.

Sanches, D. and S. Williamson (2010). Money and credit with limited commitment and theft. *Journal of Economic Theory 145*(4), 1525–1549.

Sarma, M. and J. Pais (2011). Financial inclusion and development. *Journal of international development 23*(5), 613–628.

Varian, H. R. (1992). *Microeconomic Analysis*. WW Norton.

Williamson, S. D. (1994). Liquidity and market participation. *Journal of Economic Dynamics and Control 18*(3), 629–670.

Appendix A Analytical derivations

A.1 Cost threshold $\tilde{\varepsilon}$ derivation

The cost threshold $\tilde{\varepsilon}$ satisfies

$$-\phi m_{+1} + \beta V_{+1}(m_{+1}) = -\tilde{\varepsilon} - \phi \hat{m}_{+1} + \beta \hat{V}^b_{+1}(\tilde{\varepsilon}, \hat{m}_{+1}) \tag{A.1}$$

Shifting one period backward and simplifying yields

$$\begin{aligned}\tilde{\varepsilon} = &-\phi_{-1}(\hat{m} - m) + \beta\sigma \left[u(\hat{q}_b) - u(q_b) + \hat{W}^b(\tilde{\varepsilon}, \hat{m} + \ell - p\hat{q}_b, 0, \ell) - W(\tilde{\varepsilon}, m - pq_b, 0, 0)\right] \\ &+ \beta(1-\sigma)\left[\hat{W}^b(\tilde{\varepsilon}, \hat{m} - d, d, 0) - W(\tilde{\varepsilon}, m, 0, 0)\right]\end{aligned} \tag{A.2}$$

For buyers who do not consume in the AM, we have

$$\hat{W}(\tilde{\varepsilon}, \hat{m} - d, d, 0) - W(\tilde{\varepsilon}, m, 0, 0) = (1 + i_d)\phi\hat{m} - \phi m, \tag{A.3}$$

while for buyers consuming in the AM we have

$$\hat{W}(\tilde{\varepsilon}, \hat{m} + \ell - p\hat{q}_b, 0, \ell) - W(\tilde{\varepsilon}, m - pq_b, 0, 0) = -(1 + i_\ell)\phi\ell. \tag{A.4}$$

Putting all the above together we get:

$$\tilde{\varepsilon} = -(\gamma - \beta)\phi(\hat{m} - m) + \beta\left[(1-\sigma)i_d + \sigma i_\ell\right]\phi\hat{m} + \beta\sigma\left[(u(\hat{q}_b) - (1 + i_\ell)\phi p\hat{q}_b) - (u(q_b) - \phi p q_b)\right] \tag{A.5}$$

A.2 Stationary monetary equilibrium

A stationary monetary equilibrium as defined in Definition 1 satisfies the following equilibrium conditions

$$\begin{gathered}\sigma\left[\frac{u'(q_b)}{c'(q_s)} - 1\right] = \bar{\iota}, \quad \frac{u'(\hat{q}_b)}{c'(q_s)} - 1 = i, \quad \bar{\iota} = i, \quad \phi p = c'(q_s), \quad m = pq_b, \quad \hat{m} = \sigma p\hat{q}_b, \\ q_s = \sigma\left((1 - F(\tilde{\varepsilon}))q_b + F(\tilde{\varepsilon})\hat{q}_b\right), \quad \tilde{\varepsilon} = \beta\sigma\left[(u(\hat{q}_b) - u'(\hat{q}_b)\hat{q}_b) - (u(q_b) - u'(q_b)q_b)\right], \\ M = (1 - F(\tilde{\varepsilon}))m + F(\tilde{\varepsilon})\hat{m}.\end{gathered}$$

Table 6: Buyers' consumption distribution

Population share	Consumption
$\sigma F(\tilde{\varepsilon})$	$\hat{q}_b + x^*$
$\sigma(1 - F(\tilde{\varepsilon}))$	$q_b + x^*$
$1 - \sigma$	x^*
Total	$q_s + x^*$

A.3 Derivation of the consumption Gini coefficient

The consumption Gini coefficient is twice the area under the consumption Lorenz curve which plots the cumulative share of the population against the cumulative share in total consumption. Given the distribution of consumption in Table 6, the resulting consumption Gini coefficient is

$$G = 1 - \frac{x^* + \sigma^2\left((1 - F(\tilde{\varepsilon}))^2 q_b + 2F(\tilde{\varepsilon})(1 - F(\tilde{\varepsilon}))q_b + (F(\tilde{\varepsilon}))^2 \hat{q}_b\right)}{q_s + x^*}. \tag{A.6}$$

Appendix B Proofs

B.1 Proposition 1

Proof. From (45), when $\gamma = \beta$, the interest rate prevalent in the banking sector becomes $i = 0$. Using that in (42) and (46) results in $q_b = \hat{q}_b = q_b^*$, where q_b^* is the efficient allocation that satisfies $u'(q_b^*) = c'(\sigma q_b^*)$. Replacing in (49), we get $\tilde{\varepsilon} = 0$ which implies $F(\tilde{\varepsilon}) = 0$. Using this in the market clearing condition yields $\sigma q_b = q_s$. Plugging this in the Euler equation yields the equilibrium condition

$$\frac{u'(q_b^*)}{c'(\sigma q_b^*)} = 1.$$

Under the usual assumptions on the utility function ($u'(q) > 0, u''(q) < 0, \lim_{q\to 0} u'(q) = +\infty, \lim_{q\to\infty} u'(q) = 0$) and cost function ($c'(q) > 0, c''(q) \geq 0, , \lim_{q\to 0} c'(q) = 0, \lim_{q\to\infty} c'(q) = +\infty$) we have

$$\frac{\partial \frac{u'(q_b^*)}{c'(\sigma q_b^*)}}{\partial q_b^*} < 0, \tag{B.1}$$

$$\lim_{q_b^*\to 0} \frac{u'(q_b^*)}{c'(\sigma q_b^*)} = +\infty \text{ for } \sigma \in (0, 1], \tag{B.2}$$

$$\lim_{q_b^*\to +\infty} \frac{u'(q_b^*)}{c'(\sigma q_b^*)} = 0. \tag{B.3}$$

It follows that q_b^* exists and is unique. ∎

B.2 Proposition 2

Proof. When the costs of financial access are sufficiently high, we have

$$-\phi m_{+1} + \beta V_{+1}(m_{+1}) > -\varepsilon - \phi \hat{m}_{+1} + \beta \hat{V}^b_{+1}(\varepsilon, \hat{m}_{+1}),\ \forall \varepsilon \in [\underline{\varepsilon}, \bar{\varepsilon}] \tag{B.4}$$

which implies $F(\tilde{\varepsilon}) = 0$ as $\tilde{\varepsilon} < \underline{\varepsilon}$. In this case, equilibrium conditions are

$$\sigma \left[\frac{u'(q_b)}{c'(q_s)} - 1 \right] = \bar{\iota}, \quad \phi p = c'(q_s), \quad \phi m = \phi p q_b, \quad \sigma q_b = q_s, \quad M = m. \tag{B.5}$$

Combining the consumption Euler equation and the goods market clearing, we have

$$\frac{u'(q_b)}{c'(\sigma q_b)} = 1 + \frac{\bar{\iota}}{\sigma}.$$

Under the usual assumptions on the utility function ($u'(q) > 0, u''(q) < 0, \lim_{q \to 0} u'(q) = +\infty, \lim_{q \to \infty} u'(q) = 0$) and cost function ($c'(q) > 0, c''(q) \geq 0, , \lim_{q \to 0} c'(q) = 0, \lim_{q \to \infty} c'(q) = +\infty$) we have

$$\frac{\partial \frac{u'(q_b)}{c'(\sigma q_b)}}{\partial q_b} < 0, \tag{B.6}$$

$$\lim_{q_b \to 0} \frac{u'(q_b)}{c'(\sigma q_b)} = +\infty \text{ for } \sigma \in (0, 1], \tag{B.7}$$

$$\lim_{q_b \to +\infty} \frac{u'(q_b)}{c'(\sigma q_b)} = 0. \tag{B.8}$$

It follows that q_b exists and is unique. Furthermore, since $\bar{\iota} > 0$, $q_b < q_b^*$. ■

B.3 Proposition 3

Proof. From the Euler equations (42) and (46) we can write the quantities consumed as a function of q_s

$$q_b = u'^{-1}((1 + i/\sigma)c'(q_s)), \tag{B.9}$$

$$\hat{q}_b = u'^{-1}\left((1 + i)c'(q_s)\right), \tag{B.10}$$

and use the two expressions in the market clearing condition (47) to reduce the problem to an equation in one unknown, that is

$$q_s - \sigma \left(u'^{-1}((1 + i/\sigma)c'(q_s)) + F(\tilde{\varepsilon})(u'^{-1}\left((1 + i)c'(q_s)\right) - u'^{-1}((1 + i/\sigma)c'(q_s))) \right) = 0 \tag{B.11}$$

where

$$\tilde{\varepsilon} = \beta\sigma \Bigg[u(u'^{-1}\left((1+i)c'(q_s)\right)) - (1+i)c'(q_s)u'^{-1}\left((1+i)c'(q_s)\right) \\ - u(u'^{-1}((1+i/\sigma)c'(q_s))) + (1+i/\sigma)c'(q_s)u'^{-1}((1+i/\sigma)c'(q_s)) \Bigg]. \tag{B.12}$$

We denote the above equation as $\mathcal{H}(q_s) = 0$.

We have that $\mathcal{H}(q_s)$ is continuous in q_s. In addition, it is easy to show that $\lim_{q_s \to 0} \tilde{\varepsilon} = +\infty$ which implies that $\lim_{q_s \to 0} F(\tilde{\varepsilon}) = 1$. Using this result, we can show that

$$\lim_{q_s \to 0} \mathcal{H}(q_s) = -\infty. \tag{B.13}$$

We can also show that

$$\lim_{q_s \to +\infty} \mathcal{H}(q_s) = +\infty. \tag{B.14}$$

Given the above, at least one root of $\mathcal{H}(q_s) = 0$ exists.

Finally, we show that the cost threshold $\tilde{\varepsilon} > 0$. We have

$$\tilde{\varepsilon} = \beta\sigma \left[(u(\hat{q}_b) - u'(\hat{q}_b)\hat{q}_b) - (u(q_b) - u'(q_b)q_b) \right]. \tag{B.15}$$

We know that any concave function f is bounded from above by its first-order Taylor approximation[27]

$$f(x) \leq f(y) + f'(y)(x - y). \tag{B.16}$$

Given that u is assumed to be strictly concave we can use the above to write

$$u(q_b) \leq u(\hat{q}_b) + u'(\hat{q})(q_b - \hat{q}_b) \tag{B.17}$$

and

$$u(\hat{q}_b) \leq u(q_b) + u'(q_b)(\hat{q}_b - q_b). \tag{B.18}$$

Combining these two inequalities we get

$$u'(\hat{q}_b)(\hat{q}_b - q_b) \leq u(\hat{q}_b) - u(q_b) \leq u'(q_b)(\hat{q}_b - q_b). \tag{B.19}$$

Adding $u'(q_b)q_b$ and subtracting $u'(\hat{q}_b)\hat{q}_b$ from each term and using the definition of $\tilde{\varepsilon}$, we get

[27] cf. Varian (1992, chap. 27, p. 489)

the expression

$$\sigma\beta(u'(q_b) - u'(\hat{q}_b))q_b \leq \tilde{\varepsilon} \leq \sigma\beta(u'(q_b) - u'(\hat{q}_b))\hat{q}_b. \tag{B.20}$$

This expression states that $\tilde{\varepsilon} > 0$ if $0 < q_b < \hat{q}_b$, which is true when $\gamma > \beta$ and $\sigma \in (0, 1)$. As such, it is possible to find a continuous distribution F with support $[\underline{\varepsilon}, \bar{\varepsilon}]$ such that $\tilde{\varepsilon} \in (\underline{\varepsilon}, \bar{\varepsilon})$ and $0 < F(\tilde{\varepsilon}) < 1$. One such example is the log-normal distribution which has the support $[0, +\infty]$.

Given the above, we have shown that there exists an equilibrium with limited banking participation when $\gamma > \beta$ and $\sigma \in (0, 1)$. ■

B.4 Proposition 4

Proof. When the distribution F is degenerate such that $\underline{\varepsilon} = \bar{\varepsilon} = 0$, buyers are identical since they all face no costs of accessing banks. Therefore, in equilibrium all buyers will make the same decision on whether to access the banking sector. To determine which one will be chosen we have to compare the lifetime utility of a buyer under each choice. We start first with AM consumption. Under a pure monetary equilibrium without bank access, the AM consumption solves

$$\frac{u'(q_b)}{c'(\sigma q_b)} = 1 + \frac{\bar{\iota}}{\sigma} \tag{B.21}$$

while in an economy where all agents access banks it solves

$$\frac{u'(\hat{q}_b)}{c'(\sigma \hat{q}_b)} = 1 + \bar{\iota}. \tag{B.22}$$

As long as $\sigma < 1$ and $\bar{\iota} > 0$ (i.e. $\gamma > \beta$), we have $\hat{q}_b > q_b$.

Next, we compare equilibrium production and consumption in the CM. We start first with a financially excluded buyer. With probability σ he consumed in the AM and he enters the CM with no money holdings. In this case, his labor supply is

$$h = x^* + \phi m_{+1} - T = x^* + \gamma c'(\sigma q_b)q_b - T. \tag{B.23}$$

Using

$$T = \phi(M - M_{-1}) = (\gamma - 1)\phi M_{-1} = (\gamma - 1)\frac{\phi_{-1}}{\gamma}M_{-1} = \frac{\gamma - 1}{\gamma}c'(\sigma q_b)q_b \tag{B.24}$$

in the labor supply expression above returns

$$h = x^* + (\frac{\gamma^2 - \gamma + 1}{\gamma})c'(\sigma q_b)q_b. \tag{B.25}$$

If the financially excluded buyer didn't consume in the previous AM his labor supply corresponds to

$$h = x^* + (\frac{\gamma^2 - \gamma + 1}{\gamma})c'(\sigma q_b)q_b - c'(\sigma q_b)q_b = x^* + \frac{(\gamma-1)^2}{\gamma}c'(\sigma q_b)q_b. \tag{B.26}$$

Combining the two we have the expression for the value function

$$(1-\beta)V = U(x^*) - x^* + \sigma u(q_b) - \frac{\gamma^2 + (\sigma-2)\gamma + 1}{\gamma}c'(\sigma q_b)q_b. \tag{B.27}$$

For financially included buyers who consumed in the AM, labor supply is given by

$$\hat{h} = x^* + \phi\hat{m}_{+1} + \phi(1+i)\ell - T. \tag{B.28}$$

When all buyers have access to banks the lump-sum transfer becomes

$$T = \phi(M - M_{-1}) = (\gamma-1)\phi M_{-1} = (\gamma-1)\frac{\phi_{-1}}{\gamma}M_{-1} = \frac{\gamma-1}{\gamma}\sigma c'(\sigma\hat{q}_b)\hat{q}_b. \tag{B.29}$$

Using the above and $\phi\ell = (1-\sigma)c'(\sigma\hat{q}_b)\hat{q}_b$ gives the labor supply expression

$$\hat{h} = x^* + \left(\frac{\gamma^2 - \gamma + 1}{\gamma}\sigma + (1+i)(1-\sigma)\right)c'(\sigma\hat{q}_b)\hat{q}_b. \tag{B.30}$$

For financially included buyers who didn't consume in the AM we have

$$\hat{h} = x^* + \left(\frac{\gamma^2 - \gamma + 1}{\gamma}\sigma - (1+i)\sigma\right)c'(\sigma\hat{q}_b)\hat{q}_b. \tag{B.31}$$

Combining the two we get

$$(1-\beta)\hat{V} = U(x^*) - x^* + \sigma u(\hat{q}_b) - \frac{\gamma^2 - \gamma + 1}{\gamma}\sigma c'(\sigma\hat{q}_b)\hat{q}_b. \tag{B.32}$$

By comparing the two value functions one can show that

$$\sigma u(\hat{q}_b) - \frac{\gamma^2 - \gamma + 1}{\gamma}\sigma c'(\sigma\hat{q}_b)\hat{q}_b \geq \sigma u(q_b) - \frac{\gamma^2 + (\sigma-2)\gamma + 1}{\gamma}c'(\sigma q_b)q_b \tag{B.33}$$

always holds for $\sigma \leq 1$ with a strict inequality for $\sigma < 1$. As a consequence, $\hat{V}(0,\hat{m}) > V(m)$

and all agents choose to access banks and consume $\hat{q}_b$. The equilibrium conditions reduce to

$$\frac{u'(\hat{q}_b)}{c'(q_s)} - 1 = \bar{\iota}, \quad \phi p = c'(q_s), \quad \hat{m} = \sigma p \hat{q}_b, \quad \sigma \hat{q}_b = q_s, \quad M = \hat{m}. \tag{B.34}$$

Under the usual assumptions on the utility function ($u'(q) > 0, u''(q) < 0, \lim_{q\to 0} u'(q) = +\infty, \lim_{q\to\infty} u'(q) = 0$) and cost function ($c'(q) > 0, c''(q) \geq 0, , \lim_{q\to 0} c'(q) = 0, \lim_{q\to\infty} c'(q) = +\infty$) we have

$$\frac{\partial \frac{u'(\hat{q}_b)}{c'(\sigma\hat{q}_b)}}{\partial \hat{q}_b} < 0, \tag{B.35}$$

$$\lim_{\hat{q}_b \to 0} \frac{u'(\hat{q}_b)}{c'(\sigma\hat{q}_b)} = +\infty > 0 \text{ for } \sigma \in (0,1], \tag{B.36}$$

$$\lim_{\hat{q}_b \to +\infty} \frac{u'(\hat{q}_b)}{c'(\sigma\hat{q}_b)} = 0. \tag{B.37}$$

It follows that $\hat{q}_b$ exists and is unique. Furthermore, since $\bar{\iota} > 0$, $\hat{q}_b < q_b^*$ where q_b^* is the efficient allocation that satisfies $u'(q_b^*) = c'(\sigma q_b^*)$. ■

B.5 Result 3

Proof. It is easy to show that $\frac{\hat{q}_b}{q_b}$ is increasing in γ and decreasing in σ by solving for the ratio of marginal utilities $\frac{u'(\hat{q}_b)}{u'(q_b)}$. To do that start from (42) and (46) to get

$$\frac{u'(\hat{q}_b)}{u'(q_b)} = \frac{\gamma}{\beta + \frac{\gamma-\beta}{\sigma}}. \tag{B.38}$$

It is straightforward to show that $\frac{\partial(u'(\hat{q}_b)/u'(q_b))}{\partial\gamma} < 0$ and $\frac{\partial(u'(\hat{q}_b)/u'(q_b))}{\partial\sigma} > 0$. As a result, $\frac{\partial(\hat{q}_b/q_b)}{\partial\gamma} > 0$ and $\frac{\partial(\hat{q}_b/q_b)}{\partial\sigma} < 0$. ■