Excessive Financial Intermediation in a Model with Endogenous Liquidity

Maya Eden*
World Bank

May 23, 2012

Abstract

I study a model in which the financial system is inefficiently large. Intermediation is similar, in some ways, to the private creation of money: producers can increase their liquidity at some real cost, but this is socially wasteful as it only translates into higher prices and reduces the value of liquidity. The failure of agents to internalize the effect of intermediation on the value of liquidity leads to excessive financing. In a dynamic model with heterogeneous producers and increasing intermediation costs, I show that tight regulation of the financial system is optimal.

JEL Classification: E44, G28
Keywords: Costs of the financial sector, financial intermediation, liquidity, financial regulation

*Please send comments to meden@worldbank.org. Earlier versions of this paper circulated under the titles "The Inefficiency of Financial Intermediation in General Equilibrium" and "The Wasteful 'Money Creation' Aspect of Financial Intermediation". This work is based on a chapter of my thesis at MIT. I thank Ricardo Caballero, Ben Eden, Anton Korinek, Guido Lorenzoni, Luis Serven and Ivan Werning for helpful comments. I also benefited from comments made by seminar participants at MIT and at the World Bank, and by conference participants at the Midwestern Macroeconomics conference. This paper does not necessarily reflect the views of the World Bank, its Executive Board or the countries it represents. All errors are mine.

> "Has the contribution of the modern world of finance to economic growth become so critical as to support remuneration to its participants beyond any earlier experience and expectations? Does the past profitability of and the value added by the financial industry really now justify profits amounting to as much as 35 to 40 percent of all profits by all US corporations? Can the truly enormous rise in the use of derivatives, complicated options, and highly structured financial instruments really have made a parallel contribution to economic efficiency?"

-Paul Volcker, 'The Time We Have Is Growing Short', *The New York Review of Books*, June 24th 2010.

1 Introduction

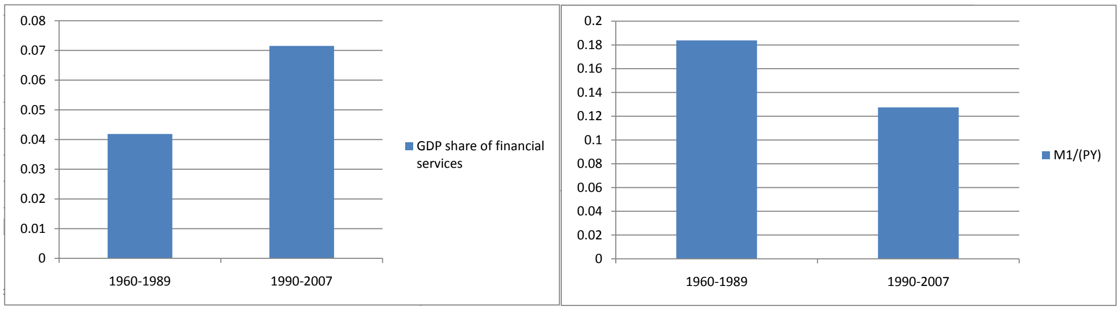
The recent financial crisis resurfaced the concern that too many productive resources are being absorbed by the financial sector, and that the real sector is too vulnerable to "mistakes" in the financial sector[1]. However, standard arguments suggest that financial intermediation improves efficiency by improving the allocation of productive resources and relaxing liquidity constraints[2]. In this paper I suggest that this may be a partial equilibrium view: while financial intermediation may appear vital when the value of liquidity is fixed, in general equilibrium, when the value of liquidity is endogenous, even the most basic form of financial intermediation may be a wasteful use of productive resources.

To develop intuition, consider the general equilibrium welfare implications of a costly technology that allows agents to produce counterfeit money. Privately, agents find it optimal to spend resources to activate this technology. However, in general equilibrium, counterfeit money merely raises the nominal price level. The equilibrium production of "useful" output (e.g., consumption goods) is lower as inputs are employed inefficiently in the creation of counterfeit money.

I argue that financial intermediation is similar, in some ways, to the production

[1] See op-eds by Friedman [2009] and Volcker [2010].

[2] See Gorton and Winton [2003] for a survey of the literature on financial intermediation, as well as Levine [2005] and McKinnon [1973] on the role of the financial sector in promoting growth and an efficient allocation of resources.



(a) GDP share of financial services (b) M1 as a fraction of nominal GDP

All figures are calculated for the US economy. The GDP share of financial services is based on data in Table 1 in Philippon [2008]. The 1960-1989 bracket represents the average GDP share of financial services in the years 1960, 1970 and 1980; the 1990-2007 bracket represents the average for the years 1990, 2000 and 2007. M1 as a fraction of nominal GDP is calculated based on data from the Federal Reserve Bank of St. Luis (FRED) database (the brackets represent average values for 1960:Q1-1989:Q4 and 1990:Q1-2007:Q4, respectively). The corresponding values for M2/PY are 0.59 for the 1960-1989 period and 0.51 for the 1990-2007 period.

Figure 1: The increase in the share of financial services in the US was accompanied by a decline in real balances as a fraction of GDP.

of counterfeit money. As a benchmark, consider an economy with identical liquidity constrained producers, and a fixed supply of production inputs. The price of production inputs is set in terms of liquidity; producers are liquidity constrained in the sense that given equilibrium prices, producers would like to hire more inputs but lack sufficient liquidity. An intermediation technology allows each producer to increase his liquidity by employing "monitoring" services. In general equilibrium, the increase in liquidity bids up the price of production inputs. Output is lower as inputs are employed inefficiently in "monitoring" services. Of course, in the presence of producer heterogeneity, the analogy to counterfeit money is incomplete because intermediation can also improve the distribution of inputs among producers. However, the improvement in the distribution of inputs need not be enough to offset the costs of financial intermediation.

The main observation of the model in this paper is that, in general equilibrium, intermediation reduces the value of liquidity[3]. This leads to excessive financial

[3]This departs slightly from standard models of financial intermediation, in which the price of inputs in terms of liquidity is essentially fixed. This view of intermediation is in the spirit of Holmstrom and Tirole [1997], in which external funds are direct inputs in investment - implicitly,

intermediation, as agents do not internalize the effect of their borrowing on the equilibrium value of liquidity. This theoretical insight applies to a wide variety of interpretations of "liquidity". In section 3, I illustrate this insight in a real model, in which liquidity is the amount of pledgeable output, and in section 4, I illustrate this insight in a model in which liquidity is money, and show that the costs of intermediation translate into a decline in real balances. Indeed, the later monetary interpretation is consistent with recent trends in the US economy. As shown in figure 1, the increase in the size of the financial system in the US has been accompanied by a decline in the amount of real balances. As the financial system grew, the amount of money shrunk relative to the price of output. This is consistent with the view that, at the aggregate level, financing reduces the equilibrium value of liquid reserves.

The effects of intermediation on the equilibrium value of liquidity is grounds for financial regulation. In a dynamic monetary model with heterogeneous producers and increasing marginal costs intermediation, I show that an unregulated financial system reduces welfare (compared to an economy with no financial system). However, I show that the welfare-maximizing reserve requirement allows for restricted levels of financial intermediation: when the costs of financial intermediation are contained, intermediation increases welfare by improving the allocation of resources.

The rest of this paper is organized as follows. In section 2, I review the related literature. In section 3, I illustrate the main mechanism using a stylized single-period model of liquidity constraints. In section 4, I generalize the conclusion of section 3 to a richer dynamic setting with money, in which the marginal cost of financial intermediation is set to be exogenously high. In section 5, I endogenize the marginal cost of intermediation by considering a stylized model of leverage externalities; I show that the welfare implications extend to this setting, and discuss the optimal reserve requirement in this context. In section 6 I conclude.

---

the price of investment goods in terms of funds is fixed at 1. This "real" view of intermediation is also in the spirit Mehra et al. [2011] and Antunes et al. [2011], who, in this framework, illustrate the efficiency of costly intermediation in dynamic macroeconomic models.

2 Related literature

This paper is related to other papers that emphasize the costs of financial intermediation. Arcand et al. [2011] show empirically that, beyond a certain threshold, there is a negative relationship between the size of the financial sector and economic growth. They present a model that attributes excessive financing to an implicit subsidy in the form of bailouts. This paper emphasizes a different mechanism through which the financial system may be overblown; in the spirit of Tobin [1984], this model highlights the difference between the private return to finance and the social return, that leads to too many productive inputs being employed on financial intermediation. The main critique in Tobin [1984] is that only a fraction of the financial system's activities can be attributed to activities that are consistent with the social objectives of finance, such as transferring funds from surplus firms to firms that need financing. Here, I argue further that even this traditional role of finance may lead to general equilibrium inefficiency through the endogenous determination of the value of liquidity.

Most closely related are Philippon [2010] and Philippon [2008], who similarly consider the macroeconomic implications of costly financial intermediation. The comparison with Philippon [2008] highlights the role of the endogenous determination of the value of liquidity: Philippon [2008] presents a real model of financial intermediation, in which monitoring allows entrepreneurs to commit future output. Despite the fact that intermediation is costly in real terms, it increases efficiency, and is socially desirable. The key difference in this paper is the endogenous determination of the value of liquidity (and the relative price of "investment" inputs), that may reverse this conclusion.

The mechanism in this paper is conceptually related to the Friedman Rule (Friedman [1969]). In a monetary economy with constant money supply, binding liquidity constraints cause consumers to waste resources on trips to the bank. The mechanism here is similar in spirit: binding liquidity constraints (on the producer's side) cause agents to spend resources inefficiently on financial intermediation. The common theme is that when agents are against a constraint, actions taken to relax the constraint may be socially inefficient.

Other papers with related mechanisms include Bolton et al. [2011] and Glode et al. [2010]. These papers emphasize the “rat race” nature of the financial sector. The idea is similar: when agents are faced with constraints, there are rents to be made from alleviating those constraints, and intermediaries inefficiently compete to extract those rents. These papers focus on the micro structure of the problem, and endogenize the costs of intermediation. Here, I take the costs of financial intermediation basically as exogenous, and focus on the dynamic macroeconomic implications of finance. An important driver of my results is the endogenous determination of the value of liquidity, a general equilibrium aspect typically ignored by this literature.

The role of the endogenous determination of the value of liquidity is similar to Bewley [1987], who considers a model in which agents with idiosyncratic income shocks hold money. In equilibrium, agents behave as if their marginal utility of holding money is constant, and use money reserves to smooth consumption. The mechanism here is similar: in the absence of financial intermediation, agents can respond to shocks by using money reserves (or liquidity). The price level is low (and the value of liquidity is high) because not all agents decide to use all of their money all of the time. In other words, “idle” money is important; the inefficiency of financial intermediation in this model stems, in part, from the reduction in “idle” money, that compromises its ability to buffer shocks.

Finally, this paper is related to the literature on the optimal regulation of the financial sector, with an emphasis on optimal reserve requirements. The literature presents different motives for regulating reserves. Farhi et al. [2009] and Cothren and Waud [1994] present models in which reserve requirements may correct market failures within the financial system. Other papers emphasize the role of reserve requirements in providing insurance and containing system risk (see, for example, Caballero and Krishnamurthy [2001] or Fernandez and Guidotti [1996]). In this paper, the role of reserve requirements is simply to reduce the extent of financial intermediation, and contain the costs absorbed by intermediation activities.

3 A simplified single period model

I begin by illustrating the intuition using a highly simplified version of the model. The model is simplified on several dimensions: it abstracts from producer heterogeneity as well as from dynamic concerns. Furthermore, the modeling of financial intermediation is extremely stylized: there will be no borrowing and lending, but rather, intermediation will amount to "monitoring services" that allow producers to pledge more of their post-production output. These simplifications are useful for developing intuition, but take away from some of the more robust conclusions of the full model developed in section 4.

Setup. Consider a single period economy with a unit measure of *producers*, indexed $i \in [0, 1]$, and a unit measure of *capital suppliers*. Capital is the only input of production. Each producer is endowed with an AK production technology, and no capital. Note that all producers share a common technology, and a common productivity parameter (A). Each capital supplier is endowed with K units of capital, but no production technology.

As producers are born without any capital, in order to produce they must purchase capital from capital suppliers. The price of capital in terms of the final good is denoted R. Capital suppliers always sell their capital at the market price (the consumption value of capital is assumed 0). The timing of the model is as follows:

1. Producers are born with a production technology (and no capital), capital suppliers are born with K units of capital (and no production technology).

2. Producers and capital suppliers trade capital for promises on post-production output.

3. Production takes place. Producers repay capital suppliers.

The amount of capital employed by producer i is denoted k_i. Producers maximize profits, which are given by production revenues minus capital costs:

$$\max_{k_i} Ak_i - Rk_i \tag{1}$$

Table 1 summarizes the notation used in this section (including notation not yet introduced).

Table 1: Section 3 notation

Notation	Variable
A	Productivity
K	Capital stock
k_i	Capital employed by producer i
l	Liquidity without monitoring
Y	Output
R	Price of capital
l'	Liquidity with monitoring
ι	Cost of intermediation

The liquidity constraint. Producers are *liquidity constrained*, in the sense that they are unable to fully commit post-production output. This friction can be seen as a reduced form formulation of a moral hazard problem, in which producers can seize some of the post-production output, and are unable to commit to refrain from doing so. For simplicity, I assume that the maximum amount of output that a producer can promise to repay is some constant l (for "liquidity")[4]:

$$Rk_i \leq l \tag{2}$$

The above constraint will be referred to as the liquidity constraint. I assume that $l < AK$, which will guarantee that the liquidity constraint binds in equilibrium.

Equilibrium without financial intermediation. To solve for the equilibrium, consider the producer's optimization problem:

$$\max_{k_i} Ak_i - Rk_i \tag{3}$$

s.t.

$$Rk_i \leq l \tag{4}$$

It is easy to show that in this framework, the liquidity constraint is binding. To see this, note that (by assumption) l is the maximum aggregate payment to

[4]Similar results can be obtained in more elaborate frameworks, in which the technology has decreasing returns and the amount of output that producers can pledge is proportional to post-production output.

capital. It follows that $R < A$:

$$RK \leq l < AK \Rightarrow R < A \tag{5}$$

As $R < A$, profits are increasing in k_i; absent the liquidity constraint, the producers would choose $k_i = \infty$. As the capital supply is finite, the liquidity constraint must bind.

The binding liquidity constraints pin down the equilibrium price of capital, R. To see this, note that each producer's demand for capital is pinned down by his liquidity constraint:

$$Rk_i = l \Rightarrow k_i = \frac{l}{R} \tag{6}$$

Capital market clearing requires that $k_i = K$ for every i. Thus, the equilibrium price of capital is given by:

$$R = \frac{l}{K} \tag{7}$$

Note that the price of capital is constrained by the aggregate supply of liquidity, and is increasing in l. Despite the fact that the liquidity constraints are binding, output is at its "first best" level, as all capital is employed in equilibrium at its most productive use:

$$Y = AK \tag{8}$$

The reason that output is unaffected by the binding liquidity constraints is that capital is supplied inelastically. In equilibrium, the entire capital stock must be employed by the producers. The binding liquidity constraints merely affect its price: the returns to capital are depressed, while producers realize positive profits.

Financial intermediation. In the economy described above, assume a financial intermediation technology that allows producers to be "monitored", thereby increasing the amount of post-production output that they can pledge from l to $l' > l$. However, monitoring is costly. For simplicity, I assume that there is a fixed cost of monitoring: a monitored producer must forgo ι units of output, which are lost on intermediation activities (in section 4, the costs of intermediation will depend the quantity of intermediated funds; in section 5 they will depend also on aggregate leverage). To summarize, the financial intermediation technology (when used) modifies the producer's problem as follows:

$$\max_{k_i} Ak_i - Rk_i - \iota \tag{9}$$

s.t.

$$Rk_i \leq l' \tag{10}$$

I will continue to assume that $l' < AK$, so that the aggregate liquidity constraint still binds.

Equilibrium with financial intermediation. For ι sufficiently small, I conjecture an equilibrium in which all producers use the financial intermediation technology. Intuitively, producers faced with binding liquidity constraints are willing to pay a real cost in order to relax their constraints.

To prove this conjecture, note that, as in the no-intermediation economy, the assumption that $l' < AK$ implies that in equilibrium, it must be the case that $R < A$. Thus, the liquidity constraint is binding for every producer.

It is left to show that given R, producers find it optimal to use the financial intermediation technology. When a producer uses the intermediation technology, he increases the amount of capital that he can employ (given R). If intermediation were free ($\iota = 0$), this would be a strict gain, as $A > R$, so profits are strictly increasing in capital. Continuity implies that producers opt for intermediation even when $\iota > 0$ but sufficiently small.

It follows that for $\iota > 0$ sufficiently small, all producers use financial intermediation in equilibrium. However, despite the fact that producers privately find intermediation profitable, it reduces aggregate output compared to the no-intermediation economy, as resources absorbed on financial intermediation are socially wasted:

$$Y = AK - \iota < AK \tag{11}$$

The intuition is straightforward: privately, each producer would like to increase his liquidity in order to increase the amount of capital that he can employ. However, as the entire capital stock is already employed, the aggregate amount of employed capital cannot increase. Intermediation works only towards increasing the price of capital, leaving producers worse off.

There is a natural analogy between financial intermediation in this model and the creation of "counterfeit money". Intermediation is a costly machine that allows producers to increase the amount of liquidity that they hold, at some real cost. In a monetary model, this would be similar to a machine that prints counterfeit money. Privately, increasing liquidity is optimal for each agent. However, similar to "counterfeit money", this activity only bids up the price of inputs. The resources spent on increasing agents' liquidity are entirely wasted from a social perspective.

The partial equilibrium view. The result that financial intermediation reduces equilibrium output may seem counter-intuitive at first. It is therefore useful to illustrate how the standard intuition can be recovered from partial equilibrium analysis of this model, that ignores the endogenous determination of the price of capital (R).

Consider a partial equilibrium analysis of the intermediation economy, in which R is fixed. The capital bill is given by:

$$RK = l' \tag{12}$$

Holding R fixed, one would erroneously conclude that financial intermediation increases the productivity of capital: absent financial intermediation, the capital bill would be bounded by $l < l'$, leaving part of the capital stock unemployed, and lowering output. More broadly, ignoring the endogenous determination of the costs of inputs leads to an overestimate of the extent to which financial intermediation increases efficiency, as, absent financial intermediation, input prices would be lower and producers would be able to employ more inputs with their internal funds.

Financial crises. While the partial equilibrium analysis leads to misleading conclusions regarding the social value of financial intermediation, it may be useful for thinking about "partial equilibrium" situations, such as financial crises. Consider a simple model of financial crises, in which the intermediation technology ceases to "work", while the price of inputs remains fixed in the short run. In other words, liquidity drops from l' to $l < l'$, but R remains at its pervious level (equation 12).

This type of financial shock would lead to a drop in employment and output, as the available liquidity is insufficient to employ all inputs at the price R. This suggests that while there is an argument for reducing the "steady-state" size of the financial system, there is still a case for bailing out the financial system during financial crises.

4 A dynamic model

The simple model in the previous section provides a general equilibrium view of financial intermediation, in which, in contrast to the partial equilibrium view, financial intermediation is purely wasteful. However, one might worry about the many stark simplifying assumptions in that model. First, the model effectively rules out inefficient uses of capital in equilibrium, as the only agents with any interest in employing capital are homogeneous "producers". In this section, this assumption will be relaxed and there will be inefficient users of capital, both with and without financial intermediation. Second, the model takes the supply of capital as exogenous, and does not allow for the supply of capital to respond to lower equilibrium prices. In the dynamic model, agents will decide how much capital to carry over to the next period. Finally, the model rules out the accumulation of liquid *reserves*. One might be concerned that in the absence of financial intermediation, producers will inefficiently accumulate liquid funds in order to finance their operations, and that such savings might come as a substitute to more productive investment in capital. In this section, I allow producers to accumulate liquid funds that can be used for the purchase of inputs. While liquidity accumulation and capital accumulation are privately substitutes, from a social perspective, they are not. It turns out that holding large amounts of liquidity is socially desirable[5].

In addition to these concerns, one might worry about the distributional implications of the model: it is fairly straightforward to show that in the simplified framework in the previous section, intermediation may be wasteful in terms of output but is not necessarily Pareto inefficient, as capital suppliers benefit from a higher price of capital. In the dynamic version of the model, this is no longer the case: intermediation reduces both equilibrium output and equilibrium consumption for every agent.

[5]This result is in the spirit of Friedman [1969].

4.1 Setup

Consider an economy with a unit measure of infinitely lived agents, indexed $i \in [0,1]$. Unlike the simplified version of the model, the decision whether to become a "producer" or a "capital supplier" will be determined endogenously. In every period, each agent receives an i.i.d productivity shock, A_i, which allows him to operate an $A_i K$ production technology. Denote by $F(\cdot)$ the cumulative density of A_i, and denote by $f(\cdot)$ the probability density function of A_i, where $f(\cdot)$ takes positive values on $[0, \bar{A}]$.

Each agent (i) has some initial endowment of capital, $k_{i,0}$, and some initial endowment of *money*, $m_{i,0}$. There is a cash in advance constraint on the purchase of existing capital.[6] Agents can buy and sell capital in the capital market at the *nominal* market price of R. In equilibrium, agents that have high productivity shocks will be buyers of capital, and agents that have low productivity shocks will be sellers of capital. The net amount of capital purchased by agent i at time t is denoted $\tilde{k}_{i,t}$. A negative $\tilde{k}_{i,t}$ indicates that the agent is a net seller of capital. Agents must pay for purchased capital in advance, and cannot sell more capital than what they have:

$$\frac{m_{i,t}}{R_t} \geq \tilde{k}_{i,t} \geq -k_{i,t} \tag{13}$$

This setup departs from the setup of the simplified model in the previous section, in that absent financial intermediation, no post-production output can be pledged. Rather, producers can only use their liquid reserves towards the purchase of capital. However, the results trivially extend to more similar settings in which producers can pledge some post-production output even without financial intermediation.

Let M denote the aggregate money supply, and let K_t denote the aggregate supply of capital at time t (K_0 is given). The money supply is assumed to be constant across time. Capital fully depreciates after one period.[7]

[6] See Abel [1985] or Stockman [1981] for models of cash in advance constraints on the purchase of inputs. However, money can be interpreted here more broadly as transferable claims on output.

[7] The analysis can be generalized to allow for a depreciation rate $1 \leq \delta < 1$, as long as $1 - \delta < \frac{\beta}{1-\beta} E(A)$ (this condition is necessary for a solution in which there is a capital market in equilibrium with $R < \infty$). In this case, the analysis carries through with a modified productivity distribution, $\hat{A} = A + 1 - \delta$. An agent that employes k units of capital can sell $A + 1 - \delta$ units of output at the end of the period.

After production takes place, agents can buy and sell the final good at the nominal price p_t. This means that producing agents can sell the final good in exchange for money, and unproductive agents can use their money holdings to buy goods. The notation used in this section and the next (including notation not yet introduced) is summarized in table 2.

Table 2: Section 3 notation

Notation	Variable
A_i	Productivity of agent i
$F(A)$	CDF of A
$f(A)$	PDF of A
$\bar{A}$	Maximum productivity
$\underline{A}$	Production cutoff
ρ	Fraction of producing agents: $1 - F(\underline{A})$
K	Capital
M	Money
Y	Output
R	Price of capital
p	Price of output
R_m	Price of liquidity
κ	Fraction of capital employed through intermediation
θ	Fraction of capital lost on intermediation
k_i	Capital endowment of i
m_i	Money endowment of i
$\tilde{k}_i$	Excess capital demand of i
$\tilde{m}_i$	Intermediation demand of i
k_i^I	Capital purchased with $\tilde{m}_i$
γ	Reserve requirement

Agents can save in two ways. First, they can carry over money to the next period. Second, they can use current output to install physical capital to be used or sold in the next period (depending on their productivity shock).

Denote the consumption of agent i at time t by $c_{i,t}$. Agents' utility function is given by:

$$U_i(\{c_{i,t}\}_{t=0}^{\infty}) = E(\sum_{t=0}^{\infty} \beta^t \ln(c_{i,t})) \tag{14}$$

Social welfare is the average of expected utilities, where all agents are weighted equally ($\int_0^1 E(U_i(\{c_{i,t}\}_{t=0}^{\infty}))di$).

4.2 The no-intermediation economy

Without financial intermediation, the timing within a period is as follows:

1. Agent i enters period t with $k_{i,t}$ units of capital, and $m_{i,t}$ units of money.

2. The productivity parameters, $A_{i,t}$, are revealed.

3. Agents exchange capital for money, at an equilibrium price of R_t (the nominal price of usable capital). Agent i can sell at most $k_{i,t}$ units of capital ($-\tilde{k}_{i,t} \leq k_{i,t}$), and purchase at most $\frac{m_{i,t}}{R_t}$ units of capital ($\tilde{k}_{i,t} \leq \frac{m_{i,t}}{R_t}$).

4. Production takes place.

5. Agents buy and sell output at a nominal price of p_t. The final good can either be used for consumption ($c_{i,t}$) or for installing capital to be used in the next period ($k_{i,t+1}$). Agents also decide how much money to carry over to the next period ($m_{i,t+1}$).

An equilibrium of the no-intermediation economy is defined as a sequence of good prices $\{p_t\}_{t=0}^{\infty}$, capital prices $\{R_t\}_{t=0}^{\infty}$, consumption sequences $\{\{c_{i,t}\}_{t=0}^{\infty}\}_{i\in[0,1]}$, capital sequences $\{\{k_{i,t}\}_{t=1}^{\infty}\}_{i\in[0,1]}$, money holding sequences $\{\{m_{i,t}\}_{t=1}^{\infty}\}_{i\in[0,1]}$ and capital purchases $\{\{\tilde{k}_{i,t}\}_{t=0}^{\infty}\}_{i\in[0,1]}$ that jointly solve the following:

1. The agent's optimization problem:

$$\max_{k_{i,t},m_{i,t},\tilde{k}_{i,t},c_{i,t}} E(\sum_{t=0}^{\infty} \beta^t \ln(c_{i,t})) \tag{15}$$

s.t.

$$p_t c_{i,t} + p_t k_{i,t+1} + m_{i,t+1} = m_{i,t} - R_t \tilde{k}_{i,t} + p_t A_{i,t}(k_{i,t} + \tilde{k}_{i,t}) \tag{16}$$

$$m_{i,t} - R_t \tilde{k}_{i,t} \geq 0 \tag{17}$$

$$\tilde{k}_{i,t} + k_{i,t} \geq 0 \tag{18}$$

$$m_{i,t+1} \geq 0 \tag{19}$$

$$k_{i,t+1} \geq 0 \tag{20}$$

$$k_{i,0} \geq 0 \text{ and } m_{i,0} \geq 0 \text{ are given} \tag{21}$$

2. Capital market clearing:

$$\int_0^1 \tilde{k}_{i,t} di = 0 \tag{22}$$

3. Money aggregation:

$$\int_0^1 m_{i,t} di = M \tag{23}$$

4. Goods market clearing:

$$\int_0^1 A_{i,t}(k_{i,t} + \tilde{k}_{i,t}) di = \int_0^1 (c_{i,t} + k_{i,t+1}) di \tag{24}$$

I restrict attention to recursive equilibria. A recursive equilibrium is defined as follows. There are three vectors of state variables, $\{A_i\}_{i\in[0,1]}$, $\{k_i\}_{i\in[0,1]}$ and $\{m_i\}_{i\in[0,1]}$. To save on notation, I will denote the set of state variables by $S = \{(A_i, k_i, m_i)\}_{i\in[0,1]}$. A recursive equilibrium of the no-intermediation economy is defined as a function $p(S)$, a function $R(S)$, a set of consumption plans $\{c_i(S)\}_{i\in[0,1]}$, capital accumulation plans $\{k_i'(S)\}_{i\in[0,1]}$, money accumulation plans $\{m_i'(S)\}_{i\in[0,1]}$ and capital purchases $\{\tilde{k}_i(S)\}_{i\in[0,1]}$ that jointly solve the following:

1. Agent i's optimization problem:

$$V(A_i, k_i, m_i, R, p) = \max_{k_i', m_i', \tilde{k}_i, c_i} \ln(c_i) + \beta E_{A_i'}(V(A_i', k_i', m_i', R', p')) \tag{25}$$

s.t.

$$pc_i + pk_i' + m_i' = m_i - R\tilde{k}_i + pA_i(k_i + \tilde{k}_i) \tag{26}$$

$$m_i - R\tilde{k}_i \geq 0 \tag{27}$$

$$\tilde{k}_i + k_i \geq 0 \tag{28}$$

$$m_i' \geq 0 \tag{29}$$

$$k_i' \geq 0 \tag{30}$$

2. Capital market clearing:
$$\int_0^1 \tilde{k}_i di = 0 \tag{31}$$

3. Money aggregation:
$$\int_0^1 m_i' di = M \tag{32}$$

4. Goods market clearing:
$$\int_0^1 A_i(k_i + \tilde{k}_i) di = \int_0^1 (c_i + k_i') di \tag{33}$$

It is fairly straightforward to show that any equilibrium is characterized by a cutoff $\underline{A}$ such that all agents with productivities $A_i > \underline{A}$ produce as much as they can (given their capital and money supplies), and all agents with productivities $A_i < \underline{A}$ do not produce. To see this, note that an agent finds it optimal to produce only if the following condition holds:

$$pA_i > R \tag{34}$$

The left hand side is the nominal return from employing one unit of capital. The right hand side is the nominal price of capital. If this condition holds, the agent can generate more revenue from employing his own capital then from selling it to another producer. The same condition also implies that the agent gains more from using his money to buy capital than from holding money: with one unit of money, the agent can buy $\frac{1}{R}$ units of capital, that generate a nominal revenue of $\frac{pA_i}{R}$. The alternative strategy of holding money yields a within-period return of 1. Dividing both sides of the above inequality by R yields the equivalent condition, $\frac{pA_i}{R} > 1$.

The above condition therefore characterizes the set of A's such that an agent with productivity A finds it optimal to produce. This set is characterized by a cutoff $\underline{A}$, since if the condition is satisfied for $\underline{A}$, it is trivially satisfied for any $A > \underline{A}$.

It will be convenient to denote the measure of agents that choose to produce in equilibrium by ρ:

$$\rho = 1 - F(\underline{A}) \tag{35}$$

The following lemma characterizes the unique[8] recursive equilibrium of the no-intermediation economy:

Lemma 1 *There is a unique recursive equilibrium in the no-intermediation economy, in which:*

1. *The real price of capital, $\frac{R}{p}$, and the production threshold, $\underline{A}$, are time invariant and jointly satisfy:*

$$\underline{A} = \frac{R}{p} = \frac{\beta \int_{\underline{A}}^{\bar{A}} A f(A) dA}{1 - \beta F(\underline{A})} \tag{36}$$

2. *The nominal price of capital R is given by:*

$$R = \frac{(1 - F(\underline{A}))M}{F(\underline{A})K} \tag{37}$$

3. *The consumption of an agent with $A_i \geq \underline{A}$, k_i and m_i is:*

$$c_i = (1 - \beta)A_i(k_i + \frac{m_i}{R}) \tag{38}$$

   The consumption of an agent with $A_i < \underline{A}$, k_i and m_i is:

$$c_i = (1 - \beta)\underline{A}(k_i + \frac{m_i}{R}) \tag{39}$$

4. *Output is given by:*

$$Y = \frac{\int_{\underline{A}}^{\bar{A}} A f(A) dA}{1 - F(\underline{A})} K \tag{40}$$

5. *Capital accumulation follows $K' = \frac{R}{p} K$.*

[8] Typically in models of fiat money, there are at least two equilibria: one in which money is valued and one in which the value of money is 0. Here, the equilibrium in which money is not valued is ruled out mechanically by the assumption that the nominal prices of capital and goods $p(S)$ and $R(S)$ are real-valued (if money were not valued, these nominal prices would be ∞).

The proof of the above lemma, together with other omitted proofs, is in the appendix. In equilibrium, there is a time-invariant cutoff, $\underline{A}$. Log utility implies that all agents consume a fraction $1-\beta$ of their wealth, and save the rest in money and capital (in equilibrium, agents are indifferent between these two saving facilities).

Agents with $A_i \geq \underline{A}$ produce as much as they can: they employ their own capital (k_i) and use their money holdings to purchase additional capital ($\tilde{k}_i = \frac{m_i}{R}$). Their wealth is their production revenue.

Agents with $A_i < \underline{A}$ do not produce. They sell their physical capital to productive agents ($\tilde{k}_i = -k_i$). Their real wealth is the sum of the revenue from their capital sales ($\frac{Rk_i}{p}$) and the value of their money holdings ($\frac{m_i}{p}$).

Note that both for producing and non-producing agents, real wealth depends on nominal prices. Producing agents care about the nominal price of capital (R), as it determines how much capital they can purchase with their liquidity ($\frac{m}{R}$). Unproductive agents are affected by the nominal price of goods, p, that determines how much output they can buy with their money holdings ($\frac{m}{p}$).

Capital accumulation is a function of the real price of capital. In other words, a low nominal price of capital does not lead to less capital accumulation, provided that the nominal price of the final good is proportionately lower as well. This is why, from a social perspective, liquid reserves (measured as $\frac{m}{R}$ or $\frac{m}{p}$) do not necessarily crowd out physical capital.

Unlike the simplified model in section 3, output is not at its "first best" level (defined as the level of output that would be produced in an economy with no liquidity constraints[9]). Capital is misallocated, since it is employed at an entire range of productivities $[\underline{A}, \bar{A}]$[10]; output would increase if capital were reallocated from relatively unproductive producing agents to more productive ones. Traditionally, financial intermediation is thought of as a remedy to this type of misallocation. However, it turns out that misallocation may be just as bad in the presence of an unregulated financial sector; in fact, if we take into account the fact that capital is employed inefficiently on financial intermediation, the allocation of capital may

[9] In this framework, the "first best" allocates the entire capital stock to the agent with the highest productivity - here, output would be $\bar{A}K$.

[10] The misallocation of capital as an equilibrium outcome of liquidity constraints is in the spirit of Kiyotaki and Moore [1997].

be even worse.

4.3 An economy with a costly intermediation technology

Consider an economy identical to the one described above, in which there is a technology that allows for financial intermediation. Agents can use a monitoring technology that allows them to pledge post-production sales. Monitored producers can borrow money from non-producing agents to finance the purchase of additional capital.

However, operating the intermediation technology is costly in real terms. With a slight departure from the model in section 3, I assume that the cost of intermediation is proportional to the amount of capital that is employed using intermediated funds. Formally, I assume that a fraction θ of each unit of capital employed through intermediation is absorbed on intermediation activities. To illustrate, consider an agent with productivity A_i that uses intermediation to finance the purchase of k_i^I units of capital. A fraction θ of each unit of capital is lost, so his net production is $A_i(1-\theta)k_i^I$.

In section 5, the value of θ will be determined endogenously in a competitive environment with aggregate leverage externalities. However, in order to highlight the inefficiency generated by the endogenous determination of the value of liquidity, it is useful to begin by carrying out the analysis while abstracting from other potential sources of inefficiency (e.g., leverage externalities).

Agents can lend both cash (m_i) and revenues from capital sales ($-R\tilde{k}_i$). A way to think about this is that there are several "rounds" (within each period) in which agents can trade capital for money, and lend the revenues from capital sales: at the beginning of the first round, agents can lend their cash and sell their capital. The revenues from capital sold in the first round can be lent; a second round opens, in which productive agents use intermediated funds to buy more capital. Revenues from second-round capital sales are lent to productive agents, who use it to buy more capital, and so on and so forth, until there is no more capital to be sold. The bottom line is that an unproductive agent's loanable funds are $m_i + Rk_i$ (assuming that all of his capital is sold, so $k_i = -\tilde{k}_i$).

Compared to the no-intermediation economy, agents have an additional choice variable which is how much money to borrow. Denote the demand for borrowing by $\tilde{m}_i$ (a negative value implies that the agent is a lender). It is assumed that agents (who are, in this model, doubling as financial firms) are subject to a reserve requirement: only a fraction $1-\gamma$ (where $0<\gamma<1$) of their portfolio can be lent, while the rest must be held in liquid reserves. Thus, the maximum amount of lending that agent i can undertake is:

$$-\tilde{m}_{i,t} \leq (1-\gamma)(m_{i,t} + R_t(-\tilde{k}_{i,t})) \tag{41}$$

The return to intermediated funds is denoted R_m. To summarize, the within-period timing of the model is modified as follows:

1. Agent i enters period t with $k_{i,t}$ units of capital, and $m_{i,t}$ units of money.

2. The productivity parameters, $A_{i,t}$, are revealed.

3. Agents exchange capital for money, at an equilibrium price of R_t (the nominal price of usable capital). Agent i can sell at most $k_{i,t}$ units of capital ($-\tilde{k}_{i,t} \leq k_{i,t}$).

   In addition, agents can borrow and lend money. However, a fraction θ of any capital purchased with borrowed money is absorbed on financial intermediation.

4. Production takes place.

5. Agents buy and sell output at a nominal price of p_t. **Agents repay R_m units of money per unit of money borrowed.** The final good can either be used for consumption ($c_{i,t}$) or for installing capital to be used in the next period ($k_{i,t+1}$). Agents also decide how much money to carry over to the next period ($m_{i,t+1}$).

Borrowers bear the costs of intermediation. An agent with productivity A will want to borrow as long as the returns to intermediated funds exceed their costs:

$$\frac{p(1-\theta)A}{R} \geq R_m \tag{42}$$

The left hand side is the nominal revenue generated by one unit of borrowed money: one unit of borrowed money can finance the purchase of $\frac{1}{R}$ units of capital. A fraction θ of each unit is lost on intermediation, so the "net" production is $\frac{(1-\theta)A}{R}$ units of output, that are sold at the nominal price p. The right hand side is the cost of one unit of borrowed money, R_m.

Note that as long as the inequality above is strict, the agent will want to borrow an infinite amount. Thus, competition among constrained producers would necessitate an equality for $A = \bar{A}$. Only agents with $A = \bar{A}$ will borrow in equilibrium, and they will be indifferent between borrowing and not borrowing. The equilibrium return to intermediation is given by:

$$R_m = \frac{p(1-\theta)\bar{A}}{R} \tag{43}$$

Given the simplifying assumption of a constant marginal cost of intermediation, the intermediation technology will be used only when it is sufficiently efficient. For unproductive agents to be willing to lend, it must be the case that the nominal return to intermediated funds is greater than the within-period nominal return on holding money, which is 1. Thus, for intermediation to take place in equilibrium it must be the case that:

$$\frac{p(1-\theta)\bar{A}}{R} = R_m \geq 1 \tag{44}$$

This condition restricts the values of θ under which the intermediation technology will be used in equilibrium. For high values of θ, the equilibrium without financial intermediation is a stable one: for example, if $\theta = 1$, no unproductive agent will be willing to lend, as the nominal return to intermediation is 0, which is lower than the return to holding money. The exact condition under which intermediation will be used in equilibrium is as follows. Denote with superscript NI equilibrium values of the no-intermediation economy. A sufficient condition for intermediation to be used in equilibrium is:

$$\frac{p^{NI}(1-\theta)\bar{A}}{R^{NI}} > 1 \tag{45}$$

Denote by θ^0 the value of θ for which the above condition holds with equality:

$$\frac{p^{NI}(1-\theta^0)\bar{A}}{R^{NI}} = 1 \tag{46}$$

If $\theta = \theta^0$, agents are indifferent between using intermediation or not; if $\theta < \theta^0$, the equilibrium without financial intermediation is not stable, as at no-intermediation equilibrium prices, unproductive agents strictly prefer lending. I will therefore focus on the parameter range $\theta < \theta^0$. In this range, there is a unique equilibrium in which unproductive agents lend as much as they can (subject to the reserve requirements).

I will skip the definition of equilibrium (which is standard) and define a recursive equilibrium directly. There are again three vectors of state variables, $\{A_i\}$, $\{k_i\}$ and $\{m_i\}$. I will continue to denote the set of state variables by $S = \{(A_i, k_i, m_i)\}_{i\in[0,1]}$. A recursive equilibrium of the intermediation economy is defined as a price function $p(S)$, a capital price function $R(S)$, a price function for intermediated funds $R_m(S)$, a set of consumption plans $\{c_i(S)\}_{i\in[0,1]}$, a set of capital accumulation plans $\{k_i'(S)\}_{i\in[0,1]}$, money accumulation plans $\{m_i'(S)\}_{i\in[0,1]}$, capital purchase plans $\{\tilde{k}_i(S)\}_{i\in[0,1]}$, intermediated money plans $\{\tilde{m}_i(S)\}_{i\in[0,1]}$ and plans for capital purchased with intermediation $\{k_i^I(S)\}_{i\in[0,1]}$ that jointly solve the following:

1. Agent i's optimization problem:

$$V(A_i, k_i, m_i, R, p) = \max_{k_i', m_i', \tilde{k}_i, \tilde{m}_i, k_i^I, c_i} \ln(c_i) + \beta E_{A_i'}(V(A_i', k_i', m_i', R', p')) \tag{47}$$

s.t.

$$pc_i + pk_i' + m_i' = \tag{48}$$

$$m_i + \tilde{m}_i(1-R_m) - R(\tilde{k}_i + k_i^I) + pA_i(k_i + \tilde{k}_i + (1-\theta)k_i^I)$$

s.t.

$$-\tilde{m}_i \leq (1-\gamma)(-R\tilde{k}_i + m_i) \tag{49}$$

$$m_i - R\tilde{k}_i \geq 0 \tag{50}$$

$$\tilde{k}_i + k_i \geq 0 \tag{51}$$

$$m_i' \geq 0 \tag{52}$$

$$k_i' \geq 0 \tag{53}$$

$$\tilde{m}_i - R k_i^I \geq 0 \tag{54}$$

2. Capital market clearing:

$$\int_0^1 (\tilde{k}_i + k_i^I) di = 0 \tag{55}$$

3. Money aggregation:

$$\int_0^1 m_i' di = M \tag{56}$$

4. Goods market clearing:

$$\int_0^1 A_i (k_i + \tilde{k}_i + (1-\theta) k_i^I) di = \int_0^1 (c_i + k_i') di \tag{57}$$

The normative properties of the equilibrium depend on θ. For θ close to 0, the "standard" intuition holds, and financial intermediation increases output. This is because the transfer of money to the most productive agents improves the allocation of capital within the productive sector. When this is done at a negligible real cost, equilibrium output increases. Intermediation may also imply a favorable redistribution of surplus, as unproductive agents can realize high returns to their intermediated funds.

However, these acclaimed benefits of financial intermediation are relevant only when the costs of intermediation are low. When the costs of intermediation are substantial, the presence of a financial sector may be socially inefficient. The following proposition summarizes this finding:

Proposition 1 *Let $\theta < \theta^0$.*

1. *There exists a unique recursive equilibrium.*

2. *For θ sufficiently large, the equilibrium is welfare inferior to the no-intermediation economy.*

3. *Assume that m_i and k_i are equally distributed across agents. Denote by $c_i(\theta)$ the equilibrium consumption of agent i given θ, and let c_i^{NI} denote the agent's consumption in the no-intermediation economy. Let ρ denote the fraction of producing agents in the no-intermediation economy ($\rho = \rho^{NI}$). Let $K(\theta)'$ denote next period's capital given θ, and let $(K')^{NI}$ denote next period's capital in the no-intermediation economy. At the limit $\theta \to \theta^0$:*

$$\lim_{\theta \to_- \theta^0} \frac{c_i^{NI} - c_i}{c_i} = \frac{(1-\rho)(1-\gamma)}{\rho^2} \tag{58}$$

$$\lim_{\theta \to_- \theta^0} K(\theta)' = (K')^{NI} \tag{59}$$

The intuition behind the welfare loss is as follows. For intermediation to improve the equilibrium allocation of resources, it must induce inefficient producers to switch over from self-financing to lending. However, if the costs of intermediation are high, self-financing producers will continue to opt for production; the funds channeled through intermediation will originate primarily from unproductive agents, who would otherwise choose to hold money. The output produced by the capital purchased with these funds would be relatively unproductive, as a large fraction is lost on intermediation.

Even when unproductive agents have only small private gains from intermediation, the presence of an intermediation technology has large effects on the equilibrium value of liquidity. As unproductive agents channel as much funds as they can to the productive sector, the nominal price of capital increases. This means that productive agents - that can realize high returns without paying the costs of intermediation - can purchase less capital with their available liquidity. Output therefore drops, as a large fraction of capital is employed through inefficient intermediation. This channel translates into lower real income for producing agents. The real wealth of non-producing agents declines as well, as a result of the decline in the value of money: the increase in the nominal price of capital lowers the real return to holding money, as the expected benefits from carrying over cash to the next period are lower. In equilibrium, this results in a higher nominal price of output, leaving unproductive agents worse off.

A corollary of Proposition 1 is that, when θ is high, the regulator can increase welfare by instituting a high reserve requirement. To see this, note that as $\gamma \to 1$, the consumption loss given by equation 58 goes to 0; as capital accumulation (at the limit) is unaffected by γ, increasing the reserve requirement increases current consumption without sacrificing future consumption. This presents a new argument for the regulation of reserves.

Partial equilibrium intuition. Similar to the single-period model, the partial-equilibrium analysis of this model leads to misleading conclusions. If we hold the nominal price of capital fixed, we will reach the inevitable conclusion that financial intermediation is essential for the economy to be in full employment, as a fraction of the capital stock is paid for with intermediated funds. In partial equilibrium, the presence of a financial sector improves the allocation of capital as it enables a larger fraction of the capital stock to be employed by the most highly productive agent, and allows for idle liquidity to be used for production purposes. However, similar to the single period model, the presence of a financial sector bids up the nominal price of capital, and worsens equilibrium capital allocation as some capital is absorbed inefficiently on financial intermediation.

Financial crises. As in the single-period model, the presence of a financial sector is not only costly in terms of output, but also a potential source of unnecessary fragility. Under the assumption that input prices are nominally sticky in the short run, an abruption in agents' ability or willingness to use financial intermediation will result in a drop in employment.

5 Endogenous intermediation costs

The previous section illustrates that the welfare implications of financial intermediation depend on the marginal cost of intermediation, θ. In this section, I augment the model to allow for the equilibrium determination of θ, and show that when the economy is unregulated, the equilibrium value of θ is in the range in which financial intermediation is welfare-reducing. This augmented model also allows for a richer discussion regarding the optimal reserve ratio and the optimal extent of financial regulation.

I assume that the costs of intermediation take the following form. Denote by κ the fraction of the capital stock that is employed through intermediation:

$$\kappa_t = \frac{\int_0^1 k_{i,t}^I di}{K_t} \tag{60}$$

The marginal cost of intermediation, θ, is assumed to be an increasing function of κ, that satisfies $\theta'(\kappa) \geq 0$, $\theta(0) = \theta'(0) = 0$, and $\theta(1) = 1$. This formalization is meant to capture two realistic features. The first is an increasing marginal cost of monitoring: a small amount of monitoring can be done relatively efficiently; however, as the extent of intermediation increases, good monitors become scarce and intermediation becomes more costly.

The second realistic feature that this model is meant to capture is the positive relationship between the the marginal cost of intermediation and aggregate leverage. There are many channels through which aggregate leverage may affect the costs of intermediation; the exact mechanism is beyond the scope of this paper. The idea is that as aggregate leverage increases, the costs of containing systemic risk increase as well (for example, through an increased probability of fire sales[11], through increased complexity[12], or increased costs of regulation and bailouts).

This stylized model of the determination of θ allows for stronger conclusions regarding the welfare implications of an unregulated intermediation sector:

Proposition 2 *Assume that there is no reserve requirement ($\gamma = 0$). There exists a unique recursive equilibrium. Denote equilibrium values of this economy with superscript I (for "intermediation"), and recall that equilibrium values of the no-intermediation economy are denoted with superscript NI. This equilibrium satisfies:*

1. *The production threshold, $\underline{A}$, is the same as in the no-intermediation economy:*

$$\underline{A}^I = \underline{A}^{NI} \tag{61}$$

$$\rho^I = \rho^{NI} \tag{62}$$

[11] See Shleifer and Vishny [2011] for a review of models of fire sales and some evidence.

[12] As in Caballero and Simsek [2010].

2. *The nominal price of capital R_t is higher than in the no-intermediation economy:*

$$R^I(S) \geq R^{NI}(S) \tag{63}$$

3. *Output is lower than in the no-intermediation economy:*

$$Y^I(S) \leq Y^{NI}(S) \tag{64}$$

4. *Capital accumulation is the same as in the no-intermediation economy:*

$$(K(S)^I)' = (K(S)^{NI})' \tag{65}$$

5. *The fraction of capital employed with intermediated funds, κ, is time invariant.*

6. *Compared to the no-intermediation economy, in the economy with financial intermediation consumption is lower for every agent in every state:*

$$c_i^I \leq c_i^{NI} \tag{66}$$

7. *In the special case in which k_i and m_i are distributed uniformly across agents, eliminating financial intermediation will increase consumption for every agent by:*

$$\frac{c_i^{NI}}{c_i^I} = \frac{1}{1-\kappa} \tag{67}$$

In equilibrium, the costs of intermediation are driven up to a point in which lenders are indifferent between lending and holding money. At this point, the return to intermediation from the perspective of lenders is 0: the gains from intermediation are exactly offset by their costs. The equilibrium determination of $\underline{A}$ is the same as in the no-intermediation economy, in which $\underline{A}$ is characterized by indifference between producing and holding liquidity at a rate of return of 1.

The fact that $\underline{A}$ remains the same as in the no-intermediation economy implies that the allocation of capital within the productive sector (excluding capital purchased with intermediated funds) remains unchanged. There are no relatively-inefficient producers that are deterred from self-financing.

What is the equilibrium productivity of capital employed through intermediation? Note that given $\gamma = 0$, indifference between self-financing and intermediation requires that:

$$R_m = \frac{p(1-\theta(\kappa))\bar{A}}{R} = \frac{p\underline{A}}{R} \Rightarrow (1-\theta(\kappa))\bar{A} = \underline{A} \tag{68}$$

Thus, capital employed with intermediated funds has the same net productivity as the least productive producer. The effect of financial intermediation on the equilibrium productivity of capital is therefore negative, as a fraction $\kappa > 0$ of the capital is employed by the least-productive technology, and the distribution of productivities within the productive sector remains unchanged.

Financial intermediation unambiguously reduces equilibrium welfare: all agents would be better off without it. In that sense, financial intermediation can be viewed as a coordination failure. Agents would like to commit to refraining from financial intermediation, and keeping its costs contained. However, when no financial intermediation is used, there is a temptation to use it: in the no-intermediation equilibrium, the price of capital is low, and productive agents can offer a high return on intermediated funds, at a low cost of intermediation. This temptation will be present as long as the private benefits from using intermediation outweigh their cost. In other words, the economy is in equilibrium only when the costs of intermediation are sufficiently high, and the nominal price of capital is sufficiently inflated.

Back-of-the-envelope calculation. Equation 67 allows for a simple quantitative answer to the question: how much better off would we be without a financial sector? The only crucial parameter is κ. To get a quantitative sense of it, I use the decline in capacity utilization in the US during the recent financial crisis. This reflects the view that the decline in capacity utilization is a result of a malfunction in the financial sector that restricted the use of intermediated funds. Under the assumption that the nominal price of capital remains fixed in the short run, the drop in intermediation translates into a drop in capacity utilization of κ.

Note that this estimate is a lower bound on κ: while the extent of intermediation severely dropped during the financial crisis, it did not entirely come to a halt. The true value of κ is likely to be higher[13]. By equation 67, this observation implies that the "true" gains from reverting to the no-intermediation equilibrium are even larger.

Between 1972 and 2010, average capacity utilization was 80.4. in 2009, this number dropped to 67.3[14]. This suggests that κ is at least:

$$\kappa \geq \frac{80.4 - 67.3}{80.4} \approx 0.16 \tag{69}$$

It follows that consumption would be at least 19 percent higher without financial intermediation:

$$\frac{c_i^{NI}}{c_i^I} = \frac{1}{1-\kappa} = \frac{1}{1-0.16} = 1.19 \tag{70}$$

This crude back-of-the-envelope calculation suggests that the elimination of the financial sector may be potentially associated with substantial welfare gains, even without taking into account the welfare implications of reducing financial fragility.[15] However, it turns out that the regulator can increase welfare beyond this level by instituting a high reserve requirement and allowing for restricted amounts of financial intermediation.

5.1 The optimal reserve requirement

The analysis above illustrates that welfare in an economy with an unregulated financial system is lower than welfare in an economy with no financial intermediation. This implies that a government facing a choice between having a financial intermediation technology and not having one should choose not to have one. But, is it possible that a sufficiently powerful government can maximize welfare

[13]Unless part of the fall in capacity utilization is due to other non-financial reasons. For example, the fall in capacity utilization can be explained in the context of a negative TFP shock, provided that capacity utilization leads to faster depreciation. In the context of such a model, the fall in capacity utilization has nothing to do with κ.

[14]Capacity utilization figures are taken from the Federal Reserve Statistical Release, Industrial Production and Capacity Utilization, December 15, 2011: http://www.federalreserve.gov/releases/g17/current.

[15]This calculation is done under the assumption that $\gamma = 0$. However, it is fairly straightforward to show that κ is decreasing in γ; thus, if the current level of κ is affected by regulation, the gains from eliminating an unregulated financial system are even higher.

by choosing some intermediate solution, in which there is some *regulated* financial intermediation?

The government's maximization problem is the following[16]:

$$\max_{\gamma\in[0,1],c_{i,t}} \int_0^1 \sum_{t=0}^{\infty} \beta^t \ln(c_{i,t}) di \tag{71}$$

s.t.

$c_{i,t}$ are equilibrium consumption sequences given the reserve ratio γ.

The regulator can trivially achieve the welfare in the no-intermediation economy by setting $\gamma = 1$: in this case, no capital is employed with intermediated funds and we have that $\kappa = 0$. This corresponds to the no-intermediation equilibrium.

The regulator can improve welfare beyond the welfare of the no-intermediation equilibrium by instituting a slightly lower reserve ratio and allowing for some intermediation to take place. This is both because it increases efficiency, and because it implies a favorable redistribution of surplus. In terms of efficiency, when θ is small, intermediation shifts resources to relatively more productive agents, at a small real cost. In terms of redistribution, restricted financial intermediation implies a high return to lending. Since, in this model, lenders are agents that got "bad draws" (low productivity shocks), this redistribution is favorable[17].

The analysis also shows that instituting a low reserve requirement is ineffective: there is a range $\gamma \in [0, \hat{\gamma}]$ in which the equilibrium is the same as in the unregulated economy. In other words, there exists a value $\hat{\gamma}$ such that the reserve requirement is a binding constraint only if $\gamma > \hat{\gamma}$.[18] This naturally implies that the optimal reserve ratio, to be effective, must satisfy $\gamma > \hat{\gamma}$.

[16] I choose not to model the costs of enforcement. It is fairly straightforward to show that the optimal reserve ratio is lower when the costs of enforcement are higher, so part of the gains from choosing a reserve ratio that is less than $\gamma = 1$ may come from saving on enforcement costs.

[17] More generally, restricted financial intermediation transfers surplus from constrained producers to liquidity suppliers (who are unconstrained in production). One could easily imagine circumstances under which this transfer is not favorable from a welfare perspective; in this case, the optimal reserve requirement is even higher than the one calculated based on this model.

[18] In the presence of enforcement costs, this implies that choosing $0 < \gamma \leq \hat{\gamma}$ is always an inferior policy choice: the equilibrium is the same as in an unregulated economy with $\gamma = 0$, so the regulator can do better by moving to $\gamma = 0$ and saving on enforcement costs.

The following proposition summarizes these results:

Proposition 3 *1. There exists a reserve ratio $\hat{\gamma} > 0$ such that for any $\gamma \leq \hat{\gamma}$, the equilibrium is the same as in the unregulated economy with $\gamma = 0$. The value of $\hat{\gamma}$ is given by:*

$$\hat{\gamma} = 1 - \frac{\rho^I \kappa^I}{(1-\rho^I)(1-\kappa^I)} \tag{72}$$

Where ρ^I is the fraction of producing agents in the unregulated equilibrium, and κ^I is the aggregate fraction of capital employed with intermediated funds in the unregulated economy.

2. The optimal reserve ratio is strictly less than 1 and strictly greater than $\hat{\gamma}$.

Back-of-the-envelope calculation. Proposition 3 allows for a quantitative estimate of the lower bound of the optimal reserve requirement, $\hat{\gamma}$. By equation 72, the value of $\hat{\gamma}$ depends on two parameters: κ^I and ρ^I. The value of κ^I is calibrated at $\kappa^I = 0.16$ (see equation 69).

To calibrate $\frac{\rho^I}{1-\rho^I}$, I use the ratio of average lending to average borrowing. Let d denote the average amount that non-producers lend (in the unregulated economy). The total amount of lending is therefore $d(1-\rho^I)$. This amounts to $\frac{d(1-\rho^I)}{\rho^I}$ units of "loans" per producer, which is also the average size of loans. It follows that:

$$\frac{\rho^I}{1-\rho^I} = \frac{\text{Average lending}}{\text{Average borrowing}} \tag{73}$$

For average borrowing, I use \$550,000, the average size of corporate and industrial loans in the US in 2007[19]. For average lending, I use \$400,000, the average size of US household wealth in 2007, excluding housing wealth[20]. Note that this is a conservative estimate, as some of household wealth may take other forms besides loans to the productive sector. However, this will tend to increase the estimate of $\hat{\gamma}$.

[19] Data source: Average Loan Size for All C&I Loans, All Commercial Banks (EAANQ), Federal Reserve Bank of St. Luis (FRED database)

[20] Data source: table 1.B, Wolff [2010].

By equation 72, $\hat{\gamma}$ is approximately given by:

$$\hat{\gamma} = 1 - \frac{\rho^I}{1-\rho^I}\frac{\kappa^I}{1-\kappa^I} \approx 1 - \frac{400,000}{550,000} \cdot \frac{0.16}{1-0.16} = 0.86 \tag{74}$$

This quantitative interpretation suggests that a *lower bound* on the optimal reserve requirement is %86. Of course, this is just a rough estimate; however, it is enough to suggest that the optimal extent of financial regulation may be higher than currently perceived.

6 Conclusion

The prevailing sense that the financial sector is "too large" may at first seem at odds with traditional economic theory. Typically, in hyper-competitive environments such as the financial sector, the profits of the sector reflect its contribution to output.

This paper suggests that the financial sector is inherently different. The profits of the financial sector do not reflect its contribution to output, because of the partial equilibrium nature of these profits. Unlike other sectors, if financial intermediation were restricted, the economy's need for it would decline as well, as producers would face lower input prices and would be better able to self-finance.

The payment to the financial sector reflects the return to intermediation from the perspective of its clients. However, in general equilibrium, financial intermediation will eliminate any arbitrage opportunity, resulting in low expected returns to intermediation, while absorbing potentially large real costs. As Friedman [2009] points out:

> "Perversely, the largest individual returns seem to flow to those whose job is to ensure that microscopically small deviations from observable regularities in asset price relationships persist for only one millisecond instead of three. These talented and energetic young citizens could surely be doing something more useful."

-Benjamin Friedman, 'Overmighty finance levies a tithe on growth'. *The Financial Times*, August 26 2009.

The inefficiency generated by the large amount of resources spent on financial intermediation calls for structural change. The model in this paper suggests that the regulator can improve welfare by instituting a high reserve requirement, that limits the extent of financial intermediation. However, there are potentially other, complementary ways to reform the financial sector. The key is to ensure that the financial system does not consume the entire surplus from the reallocation of resources.

This paper suggests that the case for financial regulation goes beyond the prevention of systemic risk. Even in the absence of any risk, there is a case for restricting financial intermediation in order to reduce the amount of resources wasted on intermediation.

References

Andrew B. Abel. Dynamic behavior of capital accumulation in a cash-in-advance model. *Journal of Monetary Economics*, 16(1):55–71, July 1985.

Antonio Antunes, Tiago Cavalcanti, and Anne Villamil. Costly intermediation and consumption smoothing. 2011.

Jean Louis Arcand, Enrico Berkes, and Ugo Panizza. Too much finance? 2011.

Truman F. Bewley. Stationary monetary equilibrium with a continuum of independently fluctuating consumers. In W. Hildenbrand and A. Mas-Colell, editors, *Contributions to Mathematical Economics in Honor of Gerard Debreu*. 1987.

Patrick Bolton, Tano Santos, and Jose A. Scheinkman. Cream skimming in financial markets. 2011.

Ricardo J. Caballero and Arvind Krishnamurthy. International and domestic collateral constraints in a model of emerging market crises. *Journal of Monetary Economics*, 48:513–548, 2001.

Ricardo J. Caballero and Alp Simsek. Fire sales in a model of complexity. 2010.

Richard D. Cothren and Roger N. Waud. On the optimality of reserve requirements. *Journal of Money, Credit and Banking*, 26(4):828–838, November 1994.

Emmanuel Farhi, Mikhail Golosov, and Aleh Tsyvinski. A theory of liquidity and regulation of financial intermediation. *Review of Economic Studies*, 76:973–992, 2009.

Roque B. Fernandez and Pablo E. Guidotti. Regulating the banking industry in transition economies: Exploring interactions between capital and reserve requirements. *The Journal of Policy Reform*, 1(1):109–134, 1996.

Benjamin Friedman. Overmighty finance levies a tithe on growth. *The Financial Times*, August 26 2009.

Milton Friedman. The optimum quantity of money. In *The Optimum Quantity of Money and Other Essays*, pages 1–50. Chicago: Aldine, 1969.

Vincent Glode, Richard Green, and Richard Lowery. Financial expertise as an arms' race. 2010.

Gary Gorton and Andrew Winton. Financial intermediation. In George Constantinides, Milt Harris, and Rene Stulz, editors, *Handbook of the Economics of Finance*, volume 1, chapter 8, pages 431–552. Elsevier, 2003.

Bengt Holmstrom and Jean Tirole. Financial intermediation, loanable funds, and the real sector. *The Quarterly Journal of Economics*, 112(3):663–691, August 1997.

Nobuhiro Kiyotaki and John Moore. Credit cycles. *Journal of Political Economy*, 105(2):211–248, 1997.

Ross Levine. Finance and growth: Theory and evidence. In Philippe Aghion and Steven Durlaf, editors, *Handbook of Economic Growth*. 2005.

Ronald I. McKinnon. *Money and Capital in Economic Development*. The Brookings Institution, 177 Massachusetts Avenue, N.W., Washington, D.C. 20036, 1973.

Rajnish Mehra, Facundo Piguillem, and Edward C. Prescott. Costly financial intermediation in neoclassical growth theory. 2011.

Thomas Philippon. The evolution of the us financial industry from 1860 to 2007: Theory and evidence. 2008.

Thomas Philippon. Financiers versus engineers: Should the financial sector be taxed or subsidized? *American Economic Journal: Macroeconomics*, 2(3):158–182, July 2010.

Andrei Shleifer and Robert Vishny. Fire sales in finance and macroeconomics. *Journal of Economic Perspectives*, 25(1):29–48, Winter 2011.

Alan C. Stockman. Anticipated inflation and the capital stock in a cash-in-advance economy. *Journal of Monetary Economics*, 8(3):387–393, 1981.

James Tobin. On the efficiency of the financial system. *Lloyds Bank Review*, (153): 1–15, July 1984.

Paul Volcker. 'the time we have is growing short'. *The New York Review of Books*, LVII(11), June 24 2010.

Edward N. Wolff. Recent trends in household wealth in the united states: Rising debt and the middle-class squeeze - an update to 2007. 2010.

A Proof of Lemma 1

The equilibrium is characterized by a cutoff $\underline{A}$, such that agents with productivity $A \geq \underline{A}$ produce as much as they can, and agents with productivity $A < \underline{A}$ do not produce. Denote $\rho = 1 - F(\underline{A})$.

First, I show that agents consume a fraction $1 - \beta$ of their wealth. To prove this, I am going to calculate $\frac{\partial EV}{\partial m}$ and $\frac{\partial EV}{\partial k}$. Note that ρ is the probability that the liquidity constraint is not binding. Let c_u denote the consumption given that the agent is unconstrained, and let c_A denote consumption in case the agent is constrained and his productivity is A.

$$\frac{\partial EV}{\partial m} = (1-\rho)\frac{1}{p}u'(c_u) + \int_{\underline{A}}^{\bar{A}} f(A)\frac{A}{R}u'(c_A)dA \tag{75}$$

$$\frac{\partial EV}{\partial k} = (1-\rho)\frac{R}{p}u'(c_u) + \int_{\underline{A}}^{\bar{A}} f(A)Au'(c_A)dA \tag{76}$$

Note that for all agents:

$$R'\frac{\partial EV'}{\partial m'} = \frac{\partial EV'}{\partial k'} \tag{77}$$

The monetary cost of an additional unit of k' is the current price of output, p. Thus, for agents to carry both money and capital into the next period, it must be the case that:

$$\frac{\partial EV'}{\partial m'} = \frac{1}{p}\frac{\partial EV'}{\partial k'} \tag{78}$$

Substituting the above,

$$\frac{\partial EV'}{\partial m'} = \frac{R'}{p}\frac{\partial EV'}{\partial m'} \Rightarrow \frac{R'}{p} = 1 \tag{79}$$

Consider the FOC of the agent's problem with respect to c, where λ is the Lagrange multiplier on the budget constraint:

$$u'(c) = p\lambda \tag{80}$$

The first order condition with respect to m' is:

$$\beta\frac{\partial EV'}{\partial m'} = \lambda \tag{81}$$

Combining the two, we get that:

$$\frac{1}{p}u'(c) = \beta\frac{\partial EV'}{\partial m'} = \beta((1-\rho')\frac{1}{p'}u'(c_u') + \int_{\underline{A}'}^{\bar{A}} f(A)\frac{A}{R'}u'(c_A')dA) \tag{82}$$

Denote the agent's end-of-the-period nominal wealth by W. W_u' denotes the agent's next period wealth if he is unconstrained, and W_A' denotes the agent's next period wealth if he is constrained with productivity A. Under the conjecture that agents consume a fraction $1-\beta$ of their wealth, we have that:

$$\frac{1}{p}\frac{1}{(1-\beta)\frac{W}{p}} = \beta((1-\rho')\frac{1}{p'}\frac{1}{(1-\beta)\frac{W'_u}{p'}} + \int_{\underline{A}'}^{\bar{A}} f(A)(\frac{A}{R'}\frac{1}{(1-\beta)\frac{W'_A}{p'}})dA) \tag{83}$$

Or:

$$\frac{1}{W} = \beta((1-\rho')\frac{1}{W'_u} + \int_{\underline{A}'}^{\bar{A}} f(A)(\frac{Ap'}{R'}\frac{1}{W'_A})dA) \tag{84}$$

If the agent is unconstrained next period, his nominal wealth is $W'_u = m' + R'k'$. If the agent is constrained next period with productivity A, his nominal wealth is $W'_A = p'A(k' + \frac{m'}{R'})$. The above condition can therefore be rewritten as:

$$\frac{1}{W} = \beta((1-\rho')\frac{1}{m'+R'k'} + \int_{\underline{A}'}^{\bar{A}} f(A)(\frac{1}{R'}\frac{1}{(k'+\frac{m'}{R'})})dA) \tag{85}$$

$$= \beta((1-\rho')\frac{1}{m'+R'k'} + \rho'\frac{1}{R'k'+m'}) = \frac{\beta}{R'k'+m'} \tag{86}$$

Or:

$$R'k' + m' = \beta W \tag{87}$$

As $R' = p$, we have that:

$$pk' + m' = \beta W \tag{88}$$

So the conjecture that the agent saves a fraction β of his income is verified.

Note the market clearing condition that determines R:

$$\rho M = (1-\rho)RK \Rightarrow R = \frac{\rho M}{(1-\rho)K} \tag{89}$$

Note that $\rho < 1$, as there are agents with 0 productivity that will sell their capital at any positive R.

By equation 79, the ratio of next periods money and capital is given by the current price level p:

$$\frac{\rho' M'}{(1-\rho')K'} = p \tag{90}$$

To finish the characterization of the equilibrium, we need to find the equilibrium p. Note that the income of the constrained agents is the entire sales revenues from production, denoted Y. The savings of constrained agents is given by:

$$\beta pY = M'_c + pK'_c \tag{91}$$

The income of unconstrained agents is the entire money supply (the money they held at the beginning of the period, and the revenues from renting out their capital). Their savings are therefore given by:

$$\beta M = M'_c + pK'_u \tag{92}$$

Combining the two, and using $M' = M$, we get that:

$$\beta(pY + M) = M + pK' \tag{93}$$

Rewriting using equation 90:

$$\beta(pY + M) = M + \frac{\rho'}{1-\rho'}M = \frac{M}{1-\rho'} \tag{94}$$

Or,

$$p = \frac{M}{Y}(\frac{1}{\beta(1-\rho')} - 1) = \frac{M}{Y}(\frac{1-\beta(1-\rho')}{\beta(1-\rho')}) \tag{95}$$

$$= \frac{M}{K\frac{\int_{\underline{A}}^{\bar{A}} Af(A)dA}{\rho}}(\frac{1-\beta(1-\rho')}{\beta(1-\rho')}) \tag{96}$$

$$p = \frac{\rho M}{K\int_{\underline{A}}^{\bar{A}} Af(A)dA}(\frac{1-\beta(1-\rho')}{\beta(1-\rho')}) \tag{97}$$

I conjecture an equilibrium with a time invariant $\underline{A}$. Under this assumption,

$$p = \frac{\rho M}{K\int_{\underline{A}}^{\bar{A}} Af(A)dA}(\frac{1-\beta(1-\rho)}{\beta(1-\rho)}) \tag{98}$$

Recall that $\underline{A}$ is characterized by the following indifference condition:

$$\frac{p\underline{A}}{R} = 1 \Rightarrow \frac{R}{p} = \underline{A} \tag{99}$$

Note that $\frac{p}{R}$ is time invariant:

$$\frac{p}{R} = \frac{\frac{\rho M}{K\int_{\underline{A}}^{\bar{A}} Af(A)dA}\left(\frac{1-\beta(1-\rho)}{\beta(1-\rho)}\right)}{\frac{\rho M}{(1-\rho)K}} = \frac{1-\beta(1-\rho)}{\beta\int_{\underline{A}}^{\bar{A}} Af(A)dA} \tag{100}$$

$$= \frac{1-\beta F(\underline{A})}{\beta\int_{\underline{A}}^{\bar{A}} Af(A)dA} \tag{101}$$

It follows that $\underline{A}$ is time invariant and equal to the solution of the following equation:

$$(1-\beta F(\underline{A}))\underline{A} = \beta\int_{\underline{A}}^{\bar{A}} Af(A)dA \tag{102}$$

Note that this equation has a solution for some $\underline{A} \in (0, \bar{A})$: For $\underline{A} = 0$, the LHS is 0 and the RHS is strictly positive. For $\underline{A} = \bar{A}$, the LHS is strictly positive and the RHS is 0. By continuity, there is an intermediate solution.

To solve for K' (and check that it is positive and less than Y), equate output supply and output demand:

$$Y = (1-\beta)(Y + \frac{M}{p}) + K' = (1-\beta)(Y + \frac{M}{\frac{M}{Y}\left(\frac{1-\beta(1-\rho)}{\beta(1-\rho)}\right)}) + K' \tag{103}$$

$$= (1-\beta)(Y + \frac{Y\beta(1-\rho)}{1-\beta(1-\rho)}) + K' = (1-\beta)\frac{Y - Y\beta(1-\rho) + Y\beta(1-\rho)}{1-\beta(1-\rho)} + K' \tag{104}$$

$$= \frac{(1-\beta)Y}{1-\beta(1-\rho)} + K' \tag{105}$$

$$\Rightarrow K' = Y(1 - \frac{1-\beta}{1-\beta(1-\rho)}) = Y\frac{1-\beta+\beta\rho-1+\beta}{1-\beta(1-\rho)} = \frac{\beta\rho}{1-\beta(1-\rho)}Y \tag{106}$$

It follows that $K' > 0$. It is easy to see that $K' < Y$, as:

$$\frac{\beta\rho}{1-\beta+\beta\rho} < \frac{\beta\rho}{\beta\rho} = 1 \tag{107}$$

To finish solving for K', substitute for Y:

$$K' = \frac{\beta\rho}{1-\beta+\beta\rho}\frac{\int_{\underline{A}}^{\bar{A}} f(A)dA}{\rho}K \tag{108}$$

$$= \frac{\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA}{1-\beta(1-\rho)}K = \frac{\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA}{1-\beta F(\underline{A})}K \tag{109}$$

From equation 100, we have the equilibrium relation:

$$K' = \frac{R}{p}K \tag{110}$$

It is left to show that this recursive equilibrium is unique. First, note that given the assumption of Ak technologies, given any R and p agent i's decision to produce depends only on A_i. Given that the distributions of k_i and m_i are orthogonal to productivity, this implies that aggregate equilibrium outcomes will depend only on aggregate state variables (K and M).

Note that the equilibrium is homogeneous with respect to the number of agents (here, normalized to 1): if we double the number of agents leaving the same distribution of money and capital (their total measure is doubled as well), the equilibrium remains unchanged from the perspective of each agent. Aggregate output and capital accumulation are simply doubled, and prices remain the same. Put differentlly, from the perspective of each agent, the equilibrium depends only on the average of the distributions of k_i and m_i.

Note also that any equilibrium has to be homogeneous in M: if the money supply is doubled for each agent, p and R should double as well, but output and capital accumulation should remain unchanged.

These two things imply that in any equilibrium K' and Y must be homogeneous of degree 1 in K, and p and R must be homogeneous of degree -1 in K. To see this, assume that the economy has $K = \lambda K_0$ units of capital. The capital endowment of agent i is k_i and the money endowment is m_i. From an aggregate perspective this is equivalent to an economy in which there is a measure λ of agents, where agent i has $\frac{k_i}{\lambda}$ units of capital and $\frac{m_i}{\lambda}$ units of money. The average capital endowment is K_0, and the average money endowment is $\frac{M}{\lambda}$. This implies that $R = \frac{R_0}{\lambda}$, $p = \frac{p_0}{\lambda}$, but $\frac{R}{p} = \frac{R_0}{p_0}$; thus, $\underline{A}$ is invariant to K. The condition that $\underline{A}$ is time invariant uniquely pins down the equilibrium described above.

B Proof of Proposition 1

Consider the partial derivatives of EV with respect to m and k. For an unconstrained agent, a unit of money yields a real return of $\gamma\frac{1}{p}+(1-\gamma)\frac{\bar{A}(1-\theta)}{R}$: a fraction γ is held as reserves, and can finance the purchase of $\frac{1}{p}$ goods. A fraction $(1-\gamma)$ can be intermediated: it can finance the purchase of $\frac{1}{R}$ units of capital, which yield a nominal return of $p\bar{A}(1-\theta)$ per unit; this nominal revenue can purchase goods at the price of p.

The derivative of EV with respect to m is therefore:

$$\frac{\partial EV}{\partial m} = F(\underline{A})(\gamma\frac{1}{p} + (1-\gamma)\frac{(1-\theta)\bar{A}}{R})u'(c_u) + \int_{\underline{A}}^{\bar{A}} f(A)\frac{A}{R}u'(c_A) \tag{111}$$

Similarly, the derivative of EV with respect to k is:

$$\frac{\partial EV}{\partial k} = F(\underline{A})(\gamma\frac{R}{p} + (1-\gamma)(1-\theta)\bar{A})u'(c_u) + \int_{\underline{A}}^{\bar{A}} f(A)Au'(c_A) \tag{112}$$

As in the no intermediation economy, since $\frac{\partial EV}{\partial k} = R\frac{\partial EV}{\partial m}$, indifference between holding money and capital requires that $R' = p$.

Similarly, agents save a fraction β of their income:

$$\frac{1}{p}u'(c) = \beta\frac{\partial EV'}{\partial m'} = \beta(F(\underline{A})(\gamma\frac{1}{p'} + (1-\gamma)\frac{(1-\theta)\bar{A}}{R'})u'(c_u') + \int_{\underline{A}}^{\bar{A}} f(A)\frac{A}{R'}u'(c_A')) \tag{113}$$

Note that given k_i and m_i, the real wealth of unconstrained agents is:

$$W_u = (\gamma\frac{1}{p} + (1-\gamma)\frac{(1-\theta)\bar{A}}{R})(m + Rk) \tag{114}$$

It is therefore possible to carry out the substitution as in the no-intermediation case, and show that agents save a fraction β of their wealth.

The nominal wage is given by the market clearing condition:

$$(1-\rho)RK = \rho M + (1-\rho)(1-\gamma)(M+RK) \Rightarrow R = \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K} \quad (115)$$

Assuming a constant ρ, this implies that:

$$R' = \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K'} = p \Rightarrow K' = \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)p} \quad (116)$$

The value of $\underline{A}$ is given by the following indifference condition:

$$\frac{p\underline{A}}{R} = \gamma + (1-\gamma)\frac{p(1-\theta)\bar{A}}{R} \quad (117)$$

The demand for savings is given by:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(K+\frac{M}{R}) + (1-\rho)(\frac{\gamma}{p} + (1-\gamma)\frac{(1-\theta)\bar{A}}{R})(M+RK)) \quad (118)$$

The supply of savings is given by:

$$K' + \frac{M}{p} = \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)p} + \frac{M}{p} = \frac{M}{p}(\frac{\rho+(1-\rho)(1-\gamma)}{\gamma(1-\rho)} + 1) \quad (119)$$

$$= \frac{M}{p}\frac{1}{\gamma(1-\rho)}$$

Equating supply and demand:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(K+\frac{M}{R}) + (1-\rho)(\frac{\gamma}{p} + (1-\gamma)\frac{(1-\theta)\bar{A}}{R})(M+RK)) \quad (120)$$

$$= \frac{M}{p}\frac{1}{\gamma(1-\rho)}$$

Multiplying by p:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(pK+\frac{pM}{R}) + (1-\rho)(\gamma + (1-\gamma)\frac{p(1-\theta)\bar{A}}{R})(M+RK)) \quad (121)$$

$$= \frac{M}{\gamma(1-\rho)}$$

Rewriting:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(pK + \frac{pM}{R}) + (1-\rho)\frac{p\underline{A}}{R}(M + RK)) = \frac{M}{\gamma(1-\rho)} \tag{122}$$

Substituting for R:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(pK + \frac{pM}{\frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K}})+ \tag{123}$$

$$(1-\rho)\frac{p\underline{A}}{\frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K}}(M + \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)})) = \frac{M}{\gamma(1-\rho)}$$

Simplifying:

$$\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(pK + \frac{p\gamma(1-\rho)K}{\rho+(1-\rho)(1-\gamma)})+ \tag{124}$$

$$(1-\rho)\frac{p\underline{A}\gamma(1-\rho)K}{M(\rho+(1-\rho)(1-\gamma))}\frac{M}{\gamma(1-\rho)}) = \frac{M}{\gamma(1-\rho)}$$

Or:

$$pK\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA(1 + \frac{\gamma(1-\rho)}{\rho+(1-\rho)(1-\gamma)})+ \tag{125}$$

$$(1-\rho)\frac{\underline{A}}{\rho+(1-\rho)(1-\gamma)}) = \frac{M}{\gamma(1-\rho)}$$

Simplifying:

$$pK\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA\frac{1}{\rho+(1-\rho)(1-\gamma)}+ \tag{126}$$

$$(1-\rho)\frac{\underline{A}}{\rho+(1-\rho)(1-\gamma)}) = \frac{M}{\gamma(1-\rho)}$$

$$\frac{pK\beta}{\rho+(1-\rho)(1-\gamma)}(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) = \frac{M}{\gamma(1-\rho)} \tag{127}$$

It follows that:

$$p = \frac{M(\rho + (1-\rho)(1-\gamma))}{\gamma(1-\rho)\beta K(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A})} \tag{128}$$

Note that $\frac{p}{R}$ is given by:

$$\frac{p}{R} = \frac{\frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)\beta K(\int_{\underline{A}}^{\bar{A}} f(A)AdA+(1-\rho)\underline{A})}}{\frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K}} = \frac{1}{\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A})} \tag{129}$$

Thus, $\underline{A}$ solves:

$$\underline{A} = \gamma\frac{R}{p} + (1-\gamma)(1-\theta)\bar{A} \Rightarrow \tag{130}$$

$$\underline{A} = \gamma\beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) + (1-\gamma)(1-\theta)\bar{A} \Rightarrow \tag{131}$$

$$\underline{A}(1-\gamma\beta F(\underline{A})) = \gamma\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\gamma)(1-\theta)\bar{A} \tag{132}$$

The above equation has a solution: for $\underline{A} = \bar{A}$, the LHS is greater than the RHS. For $\underline{A} = 0$, the LHS is smaller than the RHS. From continuity, there exists a solution.

As $\underline{A}$ is time invariant, it follows that RK is time invariant. Thus,

$$pK' = R'K' = RK \Rightarrow K' = \frac{R}{p}K \tag{133}$$

The amount of capital employed by intermediation is given by:

$$RK^I = (1-\gamma)(1-\rho)(M + RK) \Rightarrow K^I = (1-\gamma)(1-\rho)(\frac{M}{R} + K) \tag{134}$$

$$= (1-\gamma)(1-\rho)(\frac{K\gamma(1-\rho)}{\rho + (1-\rho)(1-\gamma)} + K) = K\frac{(1-\gamma)(1-\rho)}{\rho + (1-\rho)(1-\gamma)} \tag{135}$$

Let $\underline{A}(\gamma, \theta)$ denote the equilibrium value of $\underline{A}$. Note that regardless of θ, when $\gamma = 1$ the economy is the same as the no-intermediation economy. Denote by θ^0

the highest value of θ such that the no-intermediation equilibrium is no longer stable:

$$(1-\theta^0)\bar{A} = \underline{A}(\gamma = 1) \tag{136}$$

Note that for any $\theta < \theta^0$, intermediation will be used in equilibrium and the equilibrium will be as described above. For $\theta \rightarrow_- \theta^0$, $\underline{A} \rightarrow \underline{A}(\gamma = 1)$. To see this, note that for $\underline{A} = (1-\theta^0)\bar{A}$, the equation characterizing $\underline{A}$ is consistent with the no-intermediation equilibrium:

$$\underline{A}(1-\gamma\beta F(\underline{A})) = \gamma\beta \int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\gamma)(1-\theta^0)\bar{A} \tag{137}$$

Replacing $(1-\theta^0)\bar{A} = \underline{A}$, we get:

$$\underline{A}(1-\gamma\beta F(\underline{A}) - (1-\gamma)) = \gamma\beta \int_{\underline{A}}^{\bar{A}} f(A)AdA \tag{138}$$

$$\underline{A}(\gamma-\gamma\beta F(\underline{A})) = \gamma\beta \int_{\underline{A}}^{\bar{A}} f(A)AdA \tag{139}$$

This equation characterizes the equilibrium value of $\underline{A}$ in the no intermediation economy. Thus, as $\theta \rightarrow_- \theta^0$, $\underline{A}$ converges to the no intermediation economy. The real return to capital approaches the no-intermediation equilibrium as well:

$$\frac{R}{p} = \beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA + F(\underline{A})\underline{A}) = \beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA + F(\underline{A})\frac{\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA}{1-\beta F(\underline{A})}) \tag{140}$$

$$= \beta\int_{\underline{A}}^{\bar{A}} f(A)AdA(1+\frac{\beta F(\underline{A})}{1-\beta F(\underline{A})}) = \frac{\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA}{1-\beta F(\underline{A})} \tag{141}$$

Thus, the rate of capital accumulation also approaches the no-intermediation economy as $\theta \rightarrow_- \theta^0$.

However, the equilibrium prices of goods and capital are higher:

$$R = \frac{M(\rho+(1-\rho)(1-\gamma))}{\gamma(1-\rho)K} > \frac{\rho M}{\gamma(1-\rho)K} > \frac{\rho M}{(1-\rho)K} = R^{NI} \tag{142}$$

Thus, as the limit $\theta \to \theta_0$ of $\frac{R}{p}$ is the same as in the no intermediation, it can be concluded that the limiting p is higher as well.

The increase in nominal prices implies lower welfare. To see this, note that the consumption of an agent with $A_i = \bar{A}$, k_i and m_i is given by:

$$c(k_i, m_i, \bar{A}) = (1-\beta)\bar{A}(k_i + \frac{m_i}{R}) \tag{143}$$

Note that the consumption of an agent with $\bar{A} \geq A_i \geq \underline{A}$ is given by:

$$c(k_i, m_i, A_i \geq \underline{A}) = (1-\beta)A_i(k_i + \frac{m_i}{R}) = \frac{A_i}{\bar{A}}c(k_i, m_i, \bar{A}) \tag{144}$$

The consumption of an agent with k_i, m_i and $A_i < \underline{A}$ is given by:

$$c(k_i, m_i, A_i < \underline{A}) = (1-\beta)\frac{1}{p}(Rk_i + m_i) = (1-\beta)\frac{R}{p}(k_i + \frac{m_i}{R}) \tag{145}$$

$$= \frac{R}{p\bar{A}}c(k_i, m_i, \bar{A}) = \frac{\underline{A}}{\bar{A}}c(k_i, m_i, \bar{A})$$

As $\underline{A}$ is (approximately) the same with and without financial intermediation, it follows that if $c(k_i, m_i, \bar{A})$ is higher in the no-intermediation economy, so is any $c(k_i, m_i, A_i)$.

Consider the difference between $c(k_i, m_i, \bar{A})$ in the no-intermediation economy and $c(k_i, m_i, \bar{A})$ in the economy with financial intermediation (superscript NI denotes no-intermediation equilibrium values, and superscript I denotes intermediation equilibrium values):

$$c^{NI}(k_i, m_i, \bar{A}) - c^{I}(k_i, m_i, \bar{A}) = \bar{A}m_i(\frac{1}{R^{NI}} - \frac{1}{R^{I}}) \tag{146}$$

$$= \bar{A}\frac{m_i K}{M}(1-\rho)(\frac{1}{\rho} - \frac{\gamma}{\rho + (1-\rho)(1-\gamma)}) \tag{147}$$

Let's say everyone has identical endowments, $m_i = M$ and $k_i = K$:

$$c^{NI}(k_i, m_i, \bar{A}) - c^{I}(k_i, m_i, \bar{A}) = \bar{A}K(1-\rho)(\frac{1}{\rho} - \frac{\gamma}{\rho + (1-\rho)(1-\gamma)}) \tag{148}$$

In percentage terms:

$$\frac{c^{NI}(k_i, m_i, \bar{A}) - c^I(k_i, m_i, \bar{A})}{c^I(k_i, m_i, \bar{A})} = \frac{\bar{A}K(1-\rho)(\frac{1}{\rho} - \frac{\gamma}{\rho+(1-\rho)(1-\gamma)})}{\bar{A}(K + \frac{M}{R})} = \tag{149}$$

$$\frac{(1-\rho)(\frac{1}{\rho} - \frac{\gamma}{\rho+(1-\rho)(1-\gamma)})}{1 + \frac{\gamma(1-\rho)}{\rho+(1-\rho)(1-\gamma)}} = \frac{(1-\rho)(\frac{\rho+(1-\rho)(1-\gamma)}{\rho} - \gamma)}{\rho}$$

$$= \frac{(1-\rho)(\rho + (1-\rho)(1-\gamma) - \gamma\rho)}{\rho^2} = \frac{(1-\rho)(1-\gamma)}{\rho^2} > 0$$

To conclude, the consumption of each agent would be higher by a fraction of $\frac{(1-\rho)(1-\gamma)}{\rho^2}$ without financial intermediation.

As consumption is lower, and capital accumulation is the same, I conclude that welfare is lower.

C Proof of Proposition 2

The proof that agents consume a fraction $1 - \beta$ of their income, and that there is a unique equilibrium in which $\underline{A}$ is constant across time is similar to the proof of proposition 1. As above, we get that $p = R'$.

Recall that if agent i uses intermediation, it must be the case that:

$$\frac{p(1-\theta(\kappa))A_i}{R} = R_m \tag{150}$$

In equilibrium, only agents with $A_i = \bar{A}$ use intermediation. Thus,

$$\frac{p(1-\theta(\kappa))\bar{A}}{R} = R_m \tag{151}$$

Let $\underline{A}$ be such that that agents are indifferent between lending and producing:

$$\underline{A} = \frac{p(1-\theta(\kappa))\bar{A}}{R} \tag{152}$$

I conjecture an intermediate solution in which agents both lend through the financial system and hold cash. This requires that $R_m = 1$. Thus, we have that:

$$R = p\underline{A} \Rightarrow \frac{R}{p} = \underline{A} \tag{153}$$

The nominal demand for output is given by:

$$(1-\beta)(p\int_{\underline{A}}^{\bar{A}} f(A)A(K + \frac{M}{R})dA + (1-\rho)(M + RK)) + pK' \tag{154}$$

Writing everything in terms of R:

$$(1-\beta)(R\int_{\underline{A}}^{\bar{A}} f(A)A\frac{K}{\underline{A}}dA + \int_{\underline{A}}^{\bar{A}} f(A)A\frac{M}{\underline{A}}dA \tag{155}$$

$$+(1-\rho)(M + RK)) + R\frac{K'}{\underline{A}}$$

The nominal supply of output is:

$$p(\int_{\underline{A}}^{\bar{A}} f(A)A(K + \frac{M}{R})dA + (1-\theta)\bar{A}(K - \rho(K + \frac{M}{R}))) \tag{156}$$

Note that $R = p(1-\theta)\bar{A}$. Writing everything in terms of R:

$$R\int_{\underline{A}}^{\bar{A}} f(A)A\frac{K}{\underline{A}}dA + \int_{\underline{A}}^{\bar{A}} f(A)A\frac{M}{\underline{A}}dA \tag{157}$$

$$+RK(1-\rho) - \rho M$$

Equating supply and demand:

$$\beta(R\int_{\underline{A}}^{\bar{A}} f(A)A\frac{K}{\underline{A}}dA + \int_{\underline{A}}^{\bar{A}} f(A)A\frac{M}{\underline{A}}dA) = \tag{158}$$

$$(1-\beta)(1-\rho)(M + RK) + R\frac{K'}{\underline{A}} - RK(1-\rho) + \rho M$$

$$= R\frac{K'}{\underline{A}} + M((1-\beta)(1-\rho) + \rho) + RK(1-\rho)(1-\beta-1)$$

$$= R\frac{K'}{\underline{A}} + M((1-\rho) - \beta(1-\rho) + \rho) - \beta RK(1-\rho)$$

$$= R\frac{K'}{\underline{A}} + M(1-\beta(1-\rho)) - \beta RK(1-\rho)$$

Manipulating the above, we get that:

$$\beta R \int_{\underline{A}}^{\bar{A}} f(A)A\frac{K}{\underline{A}}dA - R\frac{K'}{\underline{A}} + \beta RK(1-\rho) = \tag{159}$$

$$M(1-\beta(1-\rho)) - \beta \int_{\underline{A}}^{\bar{A}} f(A)A\frac{M}{\underline{A}}dA$$

Write $K' = \lambda K$. We have that:

$$RK(\beta \int_{\underline{A}}^{\infty} f(A)A\frac{1}{\underline{A}}dA - \frac{\lambda}{\underline{A}} + \beta(1-\rho)) = \tag{160}$$

$$M(1-\beta(1-\rho) - \beta \int_{\underline{A}}^{\bar{A}} f(A)A\frac{1}{\underline{A}}dA)$$

Multiplying through by $\underline{A}$, we get that:

$$RK(\beta \int_{\underline{A}}^{\infty} f(A)AdA - \lambda + \beta(1-\rho)\underline{A}) = \tag{161}$$

$$M((1-\beta(1-\rho))\underline{A} - \beta \int_{\underline{A}}^{\bar{A}} f(A)AdA)$$

I conjecture that both sides of the equation are 0. The fact that the right hand side is equal to zero implies that:

$$(1-\beta(1-\rho))\underline{A} = \beta \int_{\underline{A}}^{\bar{A}} f(A)AdA \tag{162}$$

Or:

$$\frac{(1-\beta F(\underline{A}))\underline{A}}{\beta \int_{\underline{A}}^{\bar{A}} f(A)AdA} = 1 \tag{163}$$

Note that this equation is identical to equation 102, characterizing the production threshold in the no-intermediation economy. The production threshold in the economy with financial intermediation is therefore the same as the production threshold in the no-intermediation economy. This also means that $\frac{R}{p}$ is the same as in the no-intermediation case.

We can solve for λ based on the condition that the left hand size is equal to 0:

$$\lambda = \beta \int_{\underline{A}}^{\bar{A}} f(A) A dA + \beta(1-\rho)\underline{A} = \tag{164}$$

$$(1 - \beta(1-\rho))\underline{A} + \beta(1-\rho)\underline{A}$$

$$= \underline{A}$$

Thus, we get that both the production threshold and the evolution of physical capital are the same as in the no-intermediation economy.

The nominal price of capital is higher with financial intermediation, as intermediation bids up the price of capital. To see this, note that ρ is the same as in the no-intermediation economy, and market clearing implies:

$$\rho M = (1-\rho-\kappa) RK \Rightarrow R = \frac{\rho M}{(1-\rho-\kappa)K} = \frac{1-\rho}{1-\rho-\kappa} R^{NI} \tag{165}$$

Where R^{NI} is the equilibrium price of capital in the no-intermediation economy.

It is left to verify that there is enough liquidity to finance the equilibrium payment to capital:

$$(1-\rho)RK \le \rho M + (1-\rho)(M + RK) = M + (1-\rho)RK \tag{166}$$

This trivially holds as $M \ge 0$.

To calculate the effect on consumption, note that the consumption of an agent with $A_i = \bar{A}$, k_i and m_i is given by:

$$c(k_i, m_i, \bar{A}) = (1-\beta)\bar{A}(k_i + \frac{m_i}{R}) \tag{167}$$

Note that the consumption of an agent with $\bar{A} \ge A_i \ge \underline{A}$ is given by:

$$c(k_i, m_i, A_i \ge \underline{A}) = (1-\beta)A_i(k_i + \frac{m_i}{R}) = \frac{A_i}{\bar{A}} c(k_i, m_i, \bar{A}) \tag{168}$$

The consumption of an agent with k_i, m_i and $A_i < \underline{A}$ is given by:

$$c(k_i, m_i, A_i < \underline{A}) = (1-\beta)\frac{1}{p}(Rk_i + m_i) = (1-\beta)\frac{R}{p}(k_i + \frac{m_i}{R}) \tag{169}$$

$$= \frac{R}{p\bar{A}}c(k_i, m_i, \bar{A}) = \frac{\underline{A}}{\bar{A}}c(k_i, m_i, \bar{A})$$

As $\underline{A}$ is the same with and without financial intermediation, it follows that if $c(k_i, m_i, \bar{A})$ is higher in the no-intermediation economy, so is any $c(k_i, m_i, A_i)$.

Consider the difference between $c(k_i, m_i, \bar{A})$ in the no-intermediation economy and $c(k_i, m_i, \bar{A})$ in the economy with financial intermediation (superscript NI denotes no-intermediation equilibrium values, and superscript I denotes intermediation equilibrium values):

$$c^{NI}(k_i, m_i, \bar{A}) - c^{I}(k_i, m_i, \bar{A}) = \bar{A}m_i(\frac{1}{R^{NI}} - \frac{1}{R^{I}}) \tag{170}$$

$$= \bar{A}\frac{m_i K}{\rho M}(1 - \rho - (1 - \rho - \kappa)) = \bar{A}\frac{m_i K}{\rho M}\kappa \geq 0 \tag{171}$$

Let's say everyone has identical endowments, $m_i = M$ and $k_i = K$:

$$c^{NI}(k_i, m_i, \bar{A}) - c^{I}(k_i, m_i, \bar{A}) = \bar{A}\frac{K}{\rho}\kappa \tag{172}$$

In percentage terms:

$$\frac{\frac{K}{\rho}\kappa}{K + \frac{M}{R}} = \frac{\frac{K}{\rho}\kappa}{K(1 + \frac{1-\rho-\kappa}{\rho})} = \frac{\frac{K}{\rho}\kappa}{K\frac{\rho+1-\rho-\kappa}{\rho}} = \frac{\frac{K}{\rho}\kappa}{K\frac{1-\kappa}{\rho}} = \frac{\kappa}{1-\kappa} \tag{173}$$

To conclude:

$$\frac{c_i^{NI}(k_i, m_i, \bar{A})}{c_i^{I}(k_i, m_i, \bar{A})} = 1 + \frac{\kappa}{1-\kappa} = \frac{1}{1-\kappa} \tag{174}$$

From equations 168 and 169, it follows that for every A_i,

$$\frac{c_i^{NI}(k_i, m_i, A_i)}{c_i^{I}(k_i, m_i, A_i)} = \frac{1}{1-\kappa} \tag{175}$$

D Proof of Proposition 3

It is straightforward to show that, regardless of γ, agents save a fraction β of their income, and indifference between saving in money and in capital requires $p = R'$.

Denote the equilibrium values of the economy with an unregulated financial system with a superscript I. Note that $\hat{\gamma}$ is such that the equilibrium payment to capital with an unregulated financial system is financed by the entire liquidity in the economy:

$$(1-\rho^I)R^I K = \rho^I M + (1-\hat{\gamma})(1-\rho^I)(M + R^I K) \tag{176}$$

Note that the above has a solution for some $\gamma \in (0,1)$, as for $\gamma = 0$, the equilibrium is the equilibrium of the unregulated intermediation economy and we have that the RHS is smaller than the LHS. For $\gamma = 1$, the LHS is greater than the RHS, as the RHS is equal to the payment to capital from producing agents only (which is less than the entire payment to capital, that includes also intermediated funds).

Substituting in for R, we get that:

$$(1-\rho^I)\frac{\rho^I M}{(1-\rho^I-\kappa^I)K}K = \rho^I M + (1-\hat{\gamma})(1-\rho^I)(M + \frac{\rho^I M}{(1-\rho^I-\kappa^I)K}K) \tag{177}$$

Simplifying:

$$(1-\rho^I)\frac{\rho^I}{1-\rho^I-\kappa^I} = \rho^I + (1-\hat{\gamma})(1-\rho^I)(1 + \frac{\rho^I}{1-\rho^I-\kappa^I}) \tag{178}$$

Multiplying by $1-\rho^I-\kappa^I$:

$$(1-\rho^I)\rho^I = \rho^I(1-\rho^I) - \rho^I\kappa^I + (1-\hat{\gamma})(1-\rho^I)(1-\rho^I-\kappa^I+\rho^I) \tag{179}$$

Simplifying:

$$0 = -\rho^I\kappa^I + (1-\hat{\gamma})(1-\rho^I)(1-\kappa^I) \tag{180}$$

$$\frac{\rho^I\kappa^I}{(1-\rho^I)(1-\kappa^I)} = 1-\hat{\gamma} \tag{181}$$

As long as $\gamma < \hat{\gamma}$, there unregulated equilibrium is unchanged. Thus, a government that is constrained by $\bar{\gamma} \leq \hat{\gamma}$ will be unable to affect the equilibrium through instituting a reserve ratio.

Let R_m denote the nominal return to intermediated liquidity. If $\gamma > \hat{\gamma}$, the production threshold $\underline{A}$ is determined by the following indifference condition, stating that the marginal producer is indifferent between using his money for production and intermediation:

$$\frac{p\underline{A}}{R} = R_m \tag{182}$$

It is easy to see that the same threshold applies for the use of capital in production:

$$p\underline{A} = R_m \cdot R \tag{183}$$

The return R_m is given by:

$$R_m = \gamma + (1-\gamma)\frac{p(1-\theta)\bar{A}}{R} \tag{184}$$

Thus, we have that:

$$\frac{p\underline{A}}{R} = \gamma + (1-\gamma)\frac{p(1-\theta)\bar{A}}{R} \tag{185}$$

$$\Rightarrow (1-\gamma)p(1-\theta)\bar{A} = p\underline{A} - \gamma R \tag{186}$$

And:

$$\frac{R}{p} = \frac{1}{\gamma}(\underline{A} - (1-\gamma)(1-\theta)\bar{A}) \tag{187}$$

Note that R is given by:

$$(1-\rho)RK = \rho M + (1-\gamma)(1-\rho)(M + RK) \tag{188}$$

$$\gamma(1-\rho)RK = \rho M + (1-\gamma)(1-\rho)M \tag{189}$$

$$R = \frac{\rho M + (1-\gamma)(1-\rho)M}{\gamma(1-\rho)K} = \frac{M}{K}\frac{\rho + (1-\gamma)(1-\rho)}{\gamma(1-\rho)} \tag{190}$$

Or:

$$R = \frac{M}{K}(\frac{\rho}{\gamma(1-\rho)} + \frac{1-\gamma}{\gamma}) \tag{191}$$

Note that:

$$\frac{M}{R} = \frac{K\gamma(1-\rho)}{\rho + (1-\gamma)(1-\rho)} \tag{192}$$

To solve for p and $\underline{A}$ as a function of γ, equate nominal supply and nominal demand:

$$p(K + \frac{M}{R})(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)(1-\gamma)(1-\theta)\bar{A}) \tag{193}$$

$$= (1-\beta)(p(K + \frac{M}{R})\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)R_m(M + RK)) + pK' \tag{194}$$

Under stationarity, we have that $pK' = RK$, so the above expression can be rewritten as:

$$= (1-\beta)(K + \frac{M}{R})(p\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)R_m R) + RK \tag{195}$$

Simplifying:

$$p(K + \frac{M}{R})(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)(1-\gamma)(1-\theta)\bar{A}) \tag{196}$$

$$= (1-\beta)(K + \frac{M}{R})(1-\rho)R_m R + RK$$

Using equation 186, we have that:

$$(K + \frac{M}{R})(\beta p\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)(p\underline{A} - \gamma R)) \tag{197}$$

$$= (1-\beta)(K + \frac{M}{R})(1-\rho)\frac{p\underline{A}}{R}R + RK$$

$$= (1-\beta)(K + \frac{M}{R})(1-\rho)p\underline{A} + RK$$

Rewriting:

$$(K+\frac{M}{R})(\frac{p}{R}R(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA+(1-\rho)\underline{A})-(1-\rho)\gamma R) \tag{198}$$

$$=(1-\beta)(K+\frac{M}{R})(1-\rho)\frac{p}{R}R\underline{A}+RK$$

Dividing through by R:

$$(K+\frac{M}{R})(\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA+(1-\rho)\underline{A})-(1-\rho)\gamma) \tag{199}$$

$$=(1-\beta)(K+\frac{M}{R})(1-\rho)\frac{p}{R}\underline{A}+K$$

Note that:

$$K+\frac{M}{R}=K+\frac{K\gamma(1-\rho)}{\rho+(1-\gamma)(1-\rho)}=K\frac{\rho+(1-\gamma)(1-\rho)+\gamma(1-\rho)}{\rho+(1-\gamma)(1-\rho)} \tag{200}$$

$$=\frac{K}{\rho+(1-\gamma)(1-\rho)}$$

Rewriting the supply=demand condition:

$$\frac{K}{\rho+(1-\gamma)(1-\rho)}(\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA+(1-\rho)\underline{A})-(1-\rho)\gamma) \tag{201}$$

$$=(1-\beta)\frac{K}{\rho+(1-\gamma)(1-\rho)}(1-\rho)\frac{p}{R}\underline{A}+K$$

Dividing through by K:

$$\frac{1}{\rho+(1-\gamma)(1-\rho)}(\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA+(1-\rho)\underline{A})-(1-\rho)\gamma) \tag{202}$$

$$=(1-\beta)\frac{\underline{A}}{\rho+(1-\gamma)(1-\rho)}(1-\rho)\frac{p}{R}+1$$

$$\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) - (1-\rho)\gamma \tag{203}$$

$$= (1-\beta)\underline{A}(1-\rho)\frac{p}{R} + \rho + (1-\gamma)(1-\rho)$$

$$\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) = (1-\beta)\underline{A}(1-\rho)\frac{p}{R} + \rho + 1 - \rho \tag{204}$$

$$\frac{p}{R}(\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) = (1-\beta)\underline{A}(1-\rho)\frac{p}{R} + 1 \tag{205}$$

Multiplying by $\frac{R}{p}$:

$$\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A} = (1-\beta)\underline{A}(1-\rho) + \frac{R}{p} \tag{206}$$

Replacing $\frac{R}{p}$:

$$\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A} \tag{207}$$

$$= (1-\beta)\underline{A}(1-\rho) + \frac{1}{\gamma}(\underline{A} - (1-\gamma)(1-\theta)\bar{A})$$

For $\gamma = 1$, the solution coincides with the no-intermediation equilibrium (note that this equation with $\gamma = 1$ characterizes the equilibrium $\underline{A}$ in the no-intermediation economy).

Rewriting equation 207:

$$\beta\int_{\underline{A}}^{\bar{A}} f(A)AdA = -\beta\underline{A}(1-\rho) + \frac{1}{\gamma}(\underline{A} - (1-\gamma)(1-\theta)\bar{A}) \tag{208}$$

Note that:

$$\frac{1}{\gamma}(\underline{A} - (1-\gamma)(1-\theta)\bar{A}) = \frac{\underline{A} - (1-\theta)\bar{A}}{\gamma} + (1-\theta)\bar{A} \tag{209}$$

Recall that:

$$\underline{A} = \gamma \frac{R}{p} + (1-\gamma)(1-\theta)\bar{A} \tag{210}$$

It follows that:

$$\underline{A} - (1-\theta)\bar{A} = \gamma \frac{R}{p} - \gamma(1-\theta)\bar{A}$$

The fact that agents strictly prefer intermediation implies that:

$$\frac{p(1-\theta)\bar{A}}{R} > 1 \Rightarrow \frac{R}{p} < (1-\theta)\bar{A} \tag{211}$$

Thus, we have that:

$$\underline{A} - (1-\theta)\bar{A} < 0 \tag{212}$$

The expression in equation 209 is therefore increasing in γ. It follows that if we lower γ, the value of the expression declines. The LHS of equation 207 must decline as well. This implies that as we lower γ, the equilibrium value of $\underline{A}$ is higher.

For $\theta = 0$, output is decreasing in γ. It follows that for θ sufficiently small, there exists some $\epsilon > 0$ such that equilibrium output is higher with $\gamma = 1 - \epsilon$ than with $\gamma = 1$. Capital is employed by a more productive set of agents, and this productivity gain is not offset by an increase in intermediation costs (if θ is sufficiently small).

The distributional implications of this policy are favorable. Note that the real income of producing agents is given by:

$$I_A = A(k + \frac{m}{R}) \tag{213}$$

Thus, the real income of productive agents is declining with R. consider the derivative of R with respect to γ:

$$\frac{K}{M}\frac{\partial R}{\partial \gamma} = \frac{\partial \frac{\rho}{\gamma(1-\rho)} + \frac{1}{\gamma} - 1}{\partial \gamma} = -\frac{\rho}{(1-\rho)\gamma^2} - \frac{1}{\gamma^2} < 0 \tag{214}$$

Since R is decreasing in γ, a lower γ implies a higher R. Thus, the real income of productive agents falls. As output increases, we conclude that the real income of unproductive agents increases.

To conclude the proof, we need to show that the positive contemporaneous effects of choosing $\gamma < 1$ are not offset by a lower K'. Note that capital accumulation is given by:

$$pK' = R'K' = RK \Rightarrow K' = \frac{R}{p}K \tag{215}$$

From equation 208, $\frac{R}{p}$ is given by:

$$\frac{R}{p} = \beta(\int_{\underline{A}}^{\bar{A}} f(A)AdA + (1-\rho)\underline{A}) \tag{216}$$

This expression is increasing in $\underline{A}$, and hence, decreasing in γ. I thus conclude that K' is decreasing in γ (for $\gamma > \hat{\gamma}$).

I therefore conclude that, since $\gamma = 1 - \epsilon$ implies higher output and output growth and a better distribution of surplus than the equilibrium allocation with $\gamma = 1$, it is welfare superior to $\gamma = 1$. It follows that the optimal policy is some $\hat{\gamma} < \gamma < 1$.