Financial Development and Industrial Capital Accumulation

by

Biagio Bossone

World Bank

Summary

In a decentralized-decisions economy under uncertainty, the financial system can be seen as the complex of institutions, infrastructure, and instruments that the society adopts to minimize the costs of transacting promises under agents’ incomplete trust and limited information. Building on a microeconomic, general equilibrium model that portrays such fundamental function of finance, this study analytically shows that, in line with recent empirical evidence, the development of financial infrastructure stimulates larger and more efficient capital industrial accumulation. The study also shows that economies with more developed financial infrastructure can better absorb exogenous shocks to output. The results call for addressing a crucial question concerning financial sector reform sequencing: early in development banks provide essential financial infrastructural services as part of their exclusive relationships with borrowers, while further economic development requires such services to be provided extrinsically to the bank-borrower relationships, clearly at the expense of bank rents. Financial sector development is thus characterized by a discontinuity in that banks are to be supported early on in development, while they need to be “weakened” later on precisely to foster development. This raises the question of when and how optimally to generate and manage the discontinuity before it is forced upon the society by traumatic and costly events such as bank crises.

---

I wish to thank R. Rajan and L. Zingales for their seminal ideas on the relationship between financial and industrial development, that have largely inspired this work. Of course, I remain the only responsible for the opinions expressed in the text, which do not necessarily coincide with those of the World Bank. As usual, my deepest gratitude goes to my wife Ornella, for her unremitting support.
1. **Objective of the study**

The large body of accumulated experience with financial sector reforms worldwide and the deeper theoretical understanding of the working of financial systems show that significant economic efficiency gains are associated with financial liberalization.¹ Policymakers have come to agree that, in order to achieve sustainable and stable economic efficiency gains, financial liberalization programs must be supported by development of financial infrastructure, that is, the complex of law and legal systems, trading rules and technologies, payment and settlement systems, regulation and supervision, transparency rules and accounting practices, bankruptcy codes and contract enforcement mechanisms, that influence fundamentally the incentives to honest and prudent behavior from financial market participants.²

This study shows that in a market economy the development of financial infrastructure stimulates larger and more efficient industrial capital accumulation. The study owes a great deal to the reflection recently offered by Rajan and Zingales (1999) on the issue of financial development and industrial change. The study also shows that economies with more developed financial infrastructure can better absorb exogenous shocks to output.

The results of the study raise a crucial question concerning financial sector reform sequencing: whereas early in development banks provide essential financial infrastructural services as integral part of their exclusive relationships with borrowers, further economic development requires such services to be provided extrinsically to the bank-borrower relationships, clearly at the expense of bank rents. Financial sector development thus seems to be characterized by a fundamental discontinuity in that banks are to be supported early on in development, while later on they need to be - so to say - “weakened” precisely to foster development. This raises the question of _when_ and _how_ optimally to generate and manage the discontinuity before it is forced upon the society by traumatic and costly events such as bank crises.

In the following, section 2 briefly digresses on the concept of incomplete trust in relation to finance. Section 3 presents a model of resource allocation and asset pricing under incomplete trust. In section 4 the model is used to analyze the effect of financial infrastructural development on industrial capital accumulation, and to show the greater resilience to exogenous output shocks from an economy with a more developed financial infrastructure. Section 5 reviews recent evidence in support of the model’s predictions. The policy question concerning the discontinuity in the financial sector development process is addressed in section 6.
2. Information, trust, and finance

Considerable progress has been achieved over the last two decades in the theory of finance and financial policy by recognizing that information is intrinsically limited and asymmetrically distributed among the economic agents, and by studying the implications of such informational problems for the functioning of financial markets and institutions.

In a recent work, I have pushed the informational question further by looking at the consequences of that particular form of information inefficiency deriving from the incomplete trust characterizing economic agent relations. Incomplete trust is therein defined as the agents’ awareness that others may seek to pursue inappropriate gains either through deliberately reneging on obligations on earlier commitments, or by hiding information relevant to transactions. More in general, and taking into account that agents operate under uncertainty, the concept of trust may involve an agent’s judgement that her counterparty to a contract would make all reasonable efforts to deliver on her contract obligations.

The problem of incomplete trust raises greater complexities than that of traditional informational asymmetries in as much as it faces the agents with a form of radical uncertainty that cannot be addressed simply by providing them with more and better information on available options and incentives on possible actions. Even admitting that all agents shared the same knowledge of options and incentives on possible actions (clearly violating the principle of specialization and diversification of activities as essential prerequisites of a market economy), no knowledge would be possible of the inner motives which drive the choices of different individuals facing the same options and incentives.

As the risk of unfair exploitation of asymmetries becomes real under incomplete trust, what matters most to the agents is for them to find ways to continue to benefit from asymmetric information sets, and in general from specialized knowledge, while managing their mutual trust gaps. Information is searched by the agents to see whether and how they can trust each other, rather than for reducing their informational asymmetries. Institutions evolve to enforce compliance with contractual obligations and limit the effects of incomplete trust.

The problem of incomplete trust is crucial in financial transactions whereby anonymous agents trade current real resource claims in exchange for promises to receive back real resource claims at given points in future. Traders in promises need to ascertain whether their counterparties do their best to keep their promises, and whether they are able to use scarce information efficiently to this end. The exchange of promises requires agents to be able to rely on each other, to have means to select counterparties that they can trust with using private information efficiently and fairly, and to depend on institutions that effectively ensure contract performance.
In this light, one way to look at the function of the financial system in a decentralized-decisions economy is to consider its role to reduce transaction costs related to incomplete trust. Financial institutions and infrastructure are technologies through which the society seeks to solve the trust gaps among anonymous and distrustful agents. They specialize in revealing agent trustworthiness and asset quality, and provide agents with incentives to comply with contractual obligations.

Such a view of finance underlies the model presented in the next section.

3. The model

Developing on earlier methodologies (Bossone 1995, 1997), the model in this study derives a general form of optimal demand functions for assets traded in a multi-sector emerging economy with a bank-dominated financial environment. After describing the structure of the economy, the asset trading context is defined by: i) characterizing the asset price discounting mechanism under incomplete trust, ii) qualifying the concept of financial efficiency, and iii) formalizing the (indirect) utility content of money and financial assets. An optimal intertemporal decision framework under general equilibrium is then used to study the effects of financial system innovations on industrial capital.

3.1 The economy

The model refers to an emerging economy characterized by a bank-dominated financial system. Banks are relatively well developed - as compared to, say, an industrialized economy - while the domestic financial infrastructure is still relatively inefficient in the sense defined above. There are four agents - households, firms, banks, and non-bank financial intermediaries, one composite consumption commodity $C$ expressed in real terms, physical capital $K$ and three financial instruments expressed in nominal terms: bank deposit $D$, bank loans $L$, and equity $E$. Firms, banks, and non financial, intermediaries are owned by households.

Firms produce output $Y^0$ with given technology, sell output at price $P^C$, and turn their income to the households ($Y^h$). They produce (invest in) additional industrial technology to increase future output and finance it through bank loans and/or equity, depending on the relative cost of each form of financing. Technology $\chi_K$ combines capital $K$ and a knowledge intensity factor $\chi \in R^+$. Firms accumulate physical capital $K$ to the point where its marginal efficiency equals the marginal cost of funds

$$K = K(k(\chi K) - l(r^E - r^L) - r^E)$$

$K^*_K > 0, k^* < 0, l = L/K, L + E = K$
where $r^L$ and $r^E$ are the rates of return on loans and equity, respectively.

Households are characterized by well-behaved utility functions with regular shape throughout their domain, i.e., with $u'(\cdot)>0$ and $u''(\cdot)<0$ yielding positive risk aversion, and with $u'''(\cdot)>0$ ensuring that changes in the variance of consumption affect the agent’s expected marginal utility.

Deposits $D^0$ are issued by banks, bear nominal interest $r^D$, and are used both as means of exchange and stores of value. Banks lend deposits at interest rate $r^L$ which incorporates a positive (and assumed fixed) spread on $r^D$. Banks aim to protect the value and performance of the debt they own vis-à-vis their borrowers. In a bank-dominated environment, they establish direct relationship with firms to whom they serve as their sole, or main, source of external funds. Entry barriers due to regulation and severe information limitation require large sunk costs for new entrants. Banks specialize in knowing the borrowing firms and their business, and retain such specialized knowledge as proprietary. This keeps informational opacity in the system and raises entry barriers even further. Banks exploit their quasi-monopolistic power on the firms to ensure full contract performance. They are known by the public to possess a comparative advantage in extracting rents from the firms once these have engaged in borrowing contracts. The higher the efficiency of banks in securing rent-extraction from borrowers, the lower the transaction cost for trading deposits (see section 3.2).^5

Equities are placed with households, they are traded in the capital market by specialized non-bank financial intermediaries, and yield $r^E$ on market price $p^E$. Intermediaries select equities with the highest net return. Their earning derive from trading $E$ at discounts (premiums). Equities serve as stores of value; they may in principle be used in indirect exchange as well, but at non-zero price discounts vis-à-vis, say, bank deposits, since households know much less about individual corporations than they do of banks. The discounts involved in trading $E$ depends on the efficiency of the financial infrastructure (see section 3.2).

3.2 The asset trading context

Asset price discounting^6

Assets differ from one another in their power to convey information on their quality and the trustworthiness of their holders. Each asset is characterized by an optimal transaction time, that is, the minimum time needed to sell the asset at its best price, including as such the time it takes the buyer to assess the trustworthiness of the seller, the quality of the asset, the asset’s acceptability in indirect exchange, and the time needed to complete the transaction. Operationally, the proceeds from optimal asset sale equal the asset market price net of the minimum asset-specific (unit) transaction cost $d^{O^*}$ involved in completing the sale in the given
trading context. The shorter the time available to the holder for realizing the asset, and the lower
the asset’s acceptability in indirect exchange, the larger the extra discount on the asset market
price the holder must be willing to bear with respect to \( d^{Q^*} \) to secure the transaction. Thus, the
function of discount \( d^Q \geq d^{Q^*} \) of generic asset \( Q \) can be formalized as

\[
(2) \quad d^Q_t = d^Q (1 - \Delta^Q (\sigma_i^{-\gamma} | w_i) / \Delta^Q (\psi))
\]

where \( \Delta^Q \) is the optimal transaction time interval of \( Q \); \( \Delta^Q_0 \) is the time interval available to \( Q \)’s
holder to realize the asset; \( (\sigma_i^{-\gamma} | w_i) = \int E[\beta^{\tau_i} \sigma_i | w_i]_{\tau_i} \) reflects the agent’s expected (time
weighted) average variability of consumption from date \( t \) onward, conditional on signal \( w \) (see
section 4.2 for use of this indicator in this study), and \( \psi \in \mathbb{R}^+ \) reflects the efficiency of financial
infrastructure (see below).

Expectations of higher consumption variability increase the discount factor by shortening
\( \Delta^Q_0 \), while increases in financial efficiency - other conditions being equal - lower the discount
factors by reducing the asset optimal transaction times and, hence (ceteris paribus), the extra
discounts on suboptimal sales. For each asset,

\[
0 \leq d \leq 1
\]

\[
d = d^* \quad \text{if} \quad \Delta^Q > \Delta^Q^*
\]

\[
d \to 1 \quad \text{if} \quad \Delta^Q / \Delta^Q^* \to 0
\]

It is assumed that \( d = d^* = 0 \) if \( \Delta^Q^* = 0 \), that is, perfectly liquid assets trade at zero discount. For our
purposes, (2) can be simplified as

\[
(2b) \quad d^Q_t = d^Q (\sigma_i^{-\gamma} | w_i, \psi) \quad d^* > 0 \quad d^\psi < 0
\]

**Financial efficiency**

Financial efficiency in this model reflects the financial system’s capacity to reduce transaction
costs (and, hence, asset price discounts) associated with trading assets under incomplete trust. New
intermediation facilities and innovations in financial infrastructure lower asset optimal transaction
times and enhance safety in asset trading, by facilitating the agents in ascertaining the true quality of
assets and their counterparties to transactions. Note in this context that the concept of efficiency
implies that of safe asset trading as well.
In the bank-dominated economy of the model above, deposits are assumed to be exchanged at zero transaction costs (this assumption will be relieved later on). It is also assumed that gains in financial infrastructural efficiency reduce banks’ comparative advantage in extracting rents from borrowers, and thus lower their quasi-rents, by making information on borrower trustworthiness and asset quality more easily available in the economy: exclusive bank relationships become less valuable as contracting improves under more developed financial infrastructure. Eventually, where financial infrastructure is fully developed and the banks’ quasi-monopoly is eliminated, information is no longer concentrated in exclusive investor-borrower relationships, the value of firm portfolios become known to the market, and banks have no comparative advantage on non-bank intermediaries in extracting rents from borrowers (Rajan and Zingales, 1999). An important bearing of this is that in a bank-dominated environment banks may tend to resist innovations in financial infrastructure in an attempt to protect their franchise value.

The relationship between banking and financial infrastructural development may in fact run deeper than the necessarily simplistic assumptions used in this study, and may bear relevant political-economy implications. An argument could be that, when financial infrastructure is still in its infancy, banks and basic banking services – namely, fixed nominal debt contracts both on the liability and asset side of the intermediaries’ balance sheet – represent the optimal (if not the only possible) financial institutional response to the problem of agents’ incomplete trust. The close relationships that banks build up with their borrowers, and the banks’ tendency to make those relationships exclusive and protected, provide the natural way for making financial promises credible among distrustful agents in a world with poor law and contract enforcement mechanisms. As a result, in the early stages of economic development, banks are the financial infrastructure that bridge the trust gaps among savers and investors in the economy, and make up for much of the missing formal and impersonal mechanisms that in more developed economies are extrinsic to bank relationships and reduce the costs of financial transactions (e.g., information disclosure and accounting rules, legal and institutional arrangements for contract enforcement and investor protection, payment systems, etc.)

In fact, bank deposit and loan contracts and the nature of the banks’ relationships with depositors and borrowers are such as to minimize the information costs for creditors under conditions of severe trust incompleteness: bank depositors need only to rely on the reputation of their banks (which is easier to determine than individual borrowers’ reputation) and only ask for liquidity and safety of their deposits, while leaving to the banks the extra returns from bearing the costs and the risks from dealing directly with individual borrowers and business projects. On their side, banks select borrowers and projects and draft loan contracts in ways that reasonably assure them safe and stable returns on lending operations, while surrendering to the borrowers all eventual gains from
business extra-profits. Exclusive relationships with borrowers give banks a strong power to monitor and enforce borrowers’ compliance with obligations, even in the absence of extrinsic financial infrastructure. The benefits from this type of infrastructural type of service - in this case intrinsic in the bank-borrower relationship – are ultimately passed on to the depositors in the form of safety of their savings.

This being the case, it turns out that banks do stand a lot to lose from the development of extrinsic financial infrastructure (that is, external to them), which likely dissolves their informational quasi-monopoly, erodes their comparative advantage in rent-extraction from borrowers, and leads agents to adopt more rewarding risk-sharing financial institutions and instruments.

**Asset utility**

In the model of this study, money and financial assets act as vehicles used for transferring individual consumption decisions across time, at different speeds and power, to the point where future (contingent) consumption yields the highest expected marginal utility. Assets produce utility in terms of their power to make consumption accessible when needed. Such utility varies positively with the consumption accessible through the asset, and negatively with the cost of liquidating the asset. If an agent holds an asset for a certain length of time during which she might incur future income shocks, she can use the asset as an option to be used at any point of the holding period to avert (or limit) her consumption losses. To estimate the option’s current value, the agent conjectures the probability of having to exercise it (i.e., realize the asset) at each future date of the holding period at a given cost. This probability depends on the agent’s knowledge of the distribution of future shocks and on the agent’s use of signals to anticipate future shocks. The probability is defined as follows.

Consider a discrete and infinite time horizon $[0, \infty)$, and call $s^c = s \in S \subset \mathbb{R} \times \mathbb{R}^+$ the date-event whereby at any instant prior to $\tau$ the agent expects a consumption shock to be received at $\tau$ and mobilizes her resource endowments (that could otherwise be invested) to support consumption. Let $s^{-c}$ be the complement of $s^c$ in $S$, and let $w_t \in (0,1), \forall t < \tau$, be an appropriate transformation of current information $w_t \in \Omega_t$ (see Appendix I), where $\Omega_t$ is the information set available to the agent at $t$. Finally, consider probability space $\Theta=$\{pr($s^c = s^c_{\tau,t}$|$w_t$), $0 \leq pr(\cdot) \leq 1$ and pr($s^c = s^c_{\tau,t}$|$w_t$) + pr($s^c = s^{-c}_{\tau,t}$|$w_t$) = 1\}, wherein at every date $t$ each agent attaches a probability of occurrence to future date-events $s^c_{\tau+1}$’s, conditional on signal $w_t$.$^{10}$ The signal is such...
that the probability of occurrence of date-event \( s_{t}^{c} \) increases as \( w_{i} \) approaches one, that is,

\[
\lim_{w_{i} \to 1} [pr(s^{c} = s_{t}^{c} \mid w_{i}) = 1] = 1 \quad \text{(see Appendix I)}.
\]

Thus, the marginal utility of asset \( Q \) at \( t \), conditional on information and financial efficiency is constructed as the present value of the marginal utility from the stream of future contingent consumption accessible through the asset net of the marginal utility lost to price discounts from asset liquidation.

(3) \[ u'(Q_{t}) = u'(Q_{t} \mid w_{i} ; \psi) = \]

\[
\sum_{t=1}^{\infty} \beta^{\theta} \left[ \left( 1 - d_{t}^{Q} \left( \sigma_{t}^{\tau} \mid w_{t} , \psi \right) \right) , E[u'(P_{t}^{Q} / P_{t}^{C}) \prod_{t} R_{t}^{Q} \mid pr(s^{c} = s_{t}^{c} \mid w_{t}) / \sum_{t} pr(\cdot) \right] =
\]

\[ v(Q_{t}^{h}, p_{t}^{\tau}, \sigma_{t}^{\tau}, R_{t}^{Q} \mid w_{t}, \psi) \quad \text{with} \quad v'_{Q} < 0, v'_{p} > 0, v'_{\sigma} , v'_{R} > 0 \]

where:

the time subscript \( \theta = \tau - t \) is used for the compound interest factor

\( P_{t}^{Q} / P_{t}^{C} \) is the consumption attainable at \( \tau \) with \( Q \)-holdings valued at its \( t \)-dated price (note that expected future changes in the price of \( Q \) are incorporated in \( Q \)’s return);

\( p^{C} \) is the price of (composite) consumption commodity \( C \);

\( p_{t}^{\tau} = [E(p_{t}^{C})]_{t=1}^{\infty} \) and \( R_{t}^{Q} = [E(\prod_{t=1}^{\infty} (1 + r_{t}^{Q} - \pi_{t}))]_{t=1}^{\infty} \) are the vectors of the expected values (as of date \( t \)) of, respectively, commodity prices and compound gross real interest rates on assets.

Note that \( d^{Q} = 0 \) for perfectly liquid assets, and that for given values of \( Q \), \( r^{Q} \), \( p^{C} \), and \( \beta \), different combinations of \( d^{Q} \) and \( pr(\cdot) \) yield different values of \( v_{Q} \) (see Appendix II). In particular, the sign of the derivative with respective to output variability is indeterminate and will be discussed later on. Note also that innovations in financial efficiency increase the marginal utility of \( Q \) by reducing its discount factor. Finally, as will become more relevant in section 4, an increase in signal \( w \) reduces \( Q \)’s marginal utility by increasing both the probability and the size of suboptimal sales (which increases \( Q \)’s discount factor).
4. Financial development and industrial capital

4.1 Equilibrium resource allocation

In the exchange process, at each date the infinitely-lived $h$-th household ($h=1,\ldots,H$) uses its earnings to finance current consumption and/or to add to its stock of wealth. The household derives utility directly from current consumption and indirectly from asset holdings. With money and financial assets defined as future consumption options conditioned by transaction costs, the household maximizes at each date of its time horizon a composite utility function based on the utility delivered by current consumption and the utility produced by asset holdings.

The household orders its preferences across consumption commodities and assets based on a strictly quasi-concave, time-separable utility function $U: \mathbb{R}_+^4 \rightarrow \mathbb{R}_+^4$ defined as $U=U(C;D,E,A)$. At date $t$, the household plans its resource allocation to maximize:

$$U^H = \max_{C,D,A} \beta^0 [U(C^h_t;D^h_t,E^h_t)] \quad 0 < \beta < 1, \vartheta = \tau - t$$

subject to the intertemporal budget constraint:

$$p^t C^h_t + z^h_t \leq Y^h + r^Q_t Q^t_t + r^E_t P^E_t E^h_t$$

$$+ d^E_t \max(0, P^E_t E^h_{t-1} - P^E_t E^h_t) - d^E_t \max(0, P^E_{t-2} E^h_{t-2} - P^E_{t-1} E^h_{t-1})$$

(6) $C,D,E \geq 0$

(7) $\lim_{t \to \infty} D_t = \lim_{t \to \infty} E_t = 0$

with:

$$d^E > 0 \text{ if } \Delta t^* = \tau - (\tau - 1) = 1 > \Delta t^0 (\text{suboptimal sale}) \text{ and where:}$$

$$z^h_t \text{ is household saving and is defined as } z^h_t = L^h_t - L^h_{t-1} + \sum Q^h_t (P^Q_t Q^h_t - P^Q_{t-1} Q^h_{t-1}) \text{, and}$$

the two terms in $d^E \max(\cdot)$ represent, respectively, the gain/loss from buying/selling equity at a discount.

The solution to the plan (see Bossone, 1997) requires that at planning date $t$, for given values of $p, \sigma, \beta, \pi, r^D$ and $r^E$, and for a given signal $w$, the household selects for each date $\tau \geq t$ an allocation $(C^H_t;D^H_t,E^H_t)$ that satisfies the optimal intra-date rule
Rule (8) requires each household to equate at every instant the weighted marginal utilities derived from allocating the marginal resource unit to the available consumption commodities and assets (weighted with the inverse of their own current market price). For given expectations of future shocks to consumption, rule (8) ensures that the costs of mobilizing resources to absorb these shocks are minimized since the underlying optimization model incorporates the probability of incurring such costs (eq. 2). At each date, the prices in each market must be such that rule (8) holds across all households under the following market clearing conditions:

\[ \sum_h C^h = Y^0 = \sum_h Y^h \]
\[ \sum_h D^h = D^0 \]
\[ \sum_h E^h = E^0 \]

4.2 Innovations in financial infrastructure

By using rule (8) it is easy to show the effect of an increase in the efficiency of financial infrastructure, i.e., \( \psi^1 > \psi \), on equilibrium resource allocation and prices:

\[ v(C_i^{h*}) (p_i^{C*})^{-1} = v(D_i^{h*}, R_{\tau}^{D}) = v(E_i^{h*}, R_{\tau}^{E}) | \psi^1 (p_i^{E^*})^{-1} = \mu_i^1 < \mu_i \]

yielding a relatively higher marginal utility of equity holdings. This causes market conditions to change in

\[ \sum_h (C^h | w) < Y^0 \]
\[ \sum_h (D^h | w) < D^0 \]
\[ \sum_h (E^h | w; \psi^1) > E^0 \]

With finite (but positive) interest-elastic supply functions in all market, relations (D1)-D(3) lead to a new equilibrium with a larger stock and higher price of equity (i.e., a lower rate of return on
equities), lower current consumption and lower commodity price, lower deposits and a higher interest rate on deposits. Portfolio adjustments across all households re-establish rule (8). As a result, the share of (deposit-backed) bank loans to capital decreases and equity finance increases. Note, however, that as the increase in the equilibrium stock of $E$ is greater in absolute value than the drop in deposit holdings (since the new equilibrium requires a reduction in current consumption), there is a net increase in overall investment financing so that the economy’s capital stock is larger after the innovation in financial infrastructure.

This result holds also if the increase in financial infrastructural efficiency has a positive effect on bank deposits as well. The assumption of a zero transaction cost of deposits may in fact be removed by acknowledging that there is room for innovations, say, in the payment system to strengthen agents’ trust in the exchange of deposits and lower deposits’ optimal transaction time. If this occurs, two cases can be considered: in the first, since the economy portrayed is bank-dominated and banks are assumed to be at a relatively advanced stage of development, the marginal benefit from higher financial efficiency is lower for bank intermediation than for non-bank intermediation (which amounts to assuming convexity of $d_\psi^*$); in other words, higher financial efficiency reduces the transactions costs of equities more than those of deposits. Under this assumption, the result above holds although the overall shift from loans to equity finance is smaller than in the standard case above.

In the second (more extreme) case, the marginal benefit from higher financial efficiency is the same for both bank and non-bank intermediation. Here, there is no substitution of equities for loans; yet, both deposit and equity holdings equally increase as current consumption decreases and the equilibrium returns on both decrease. This follows from the marginal utility of deposit and equities increasing vis-à-vis that of current consumption. Thus, in all cases investment finance grows and, other things equals, the economy’s equilibrium capital stock is larger than in the absence of financial infrastructure innovation.  

The substitution of loans for equity, however, has an important bearing on the quality of industrial capital. At lower levels of financial development where, ceteris paribus, asset price discounts are higher, industrial capital assets can be financed more easily the lesser their knowledge-intensive factor $\chi$ : traditional, more straightforward, and easier-to-understand technologies are preferred by risk-averse investors as they are more liquid. This seems to be especially the case in bank-dominated systems, where banks’ exclusive knowledge of borrowers and their business makes bank assets illiquid in the event of borrower default (Diamond and Rajan, 1999). To reduce liquidity risk, banks require borrowers to supply large collateral in the form of physical capital, and allocate
relatively more funds to traditional technologies that are easier to re-sell to new investors in the event of failure of the original borrowers.

Conversely, with more developed financial infrastructure, where information is not concentrated in exclusive investor-borrower relationships and both entrepreneurs’ trustworthiness and assets’ quality can be assessed at lower costs, for every dollar worth of asset price discount the technologies underlying financial claims on capital should be expected to incorporate higher knowledge intensity. Equivalently, the same quantum of knowledge incorporated in a unit of capital should sell at lower discounts in economy with more developed financial infrastructure. Innovations in financial infrastructure should therefore increase the marginal efficiency of capital $k = k(\chi K)$ at each level of physical capital stock.

Under fairly general assumptions, the model thus predicts that, as the financial system develops, equity finance should increase relatively more than bank loans, the stock of physical capital should expand, less physical capital should be immobilized in collateral against bank loans, and the economy’s industrial capital should evolve toward incorporating more sophisticated technologies.

The increase in knowledge-intensity of industrial capital can alternatively be explained by the shift in types of financial contracts following financial infrastructure innovation and its effect on asset prices. The fixed nominal debt embedded in bank deposit contacts lead banks to match their liabilities with a stable nominal value of their assets. This they do by predominantly financing safer, more traditional, and largely collateralized investments, at the expense of potential extra-rents from riskier projects that, in any case, would not be captured by the banks through their standard credit contracts. As the decline in asset price discounts and equilibrium returns following financial infrastructure innovation diverts resources from bank deposits to equity holdings, non-bank intermediaries may more freely seek higher earnings from investments in more knowledge-intensive, less collateralized projects.
Note that the two explanations provided above are not mutually exclusive. The results of an increase in financial infrastructure efficiency in terms of more and better industrial capital are depicted in fig.1, where schedule $E$ is the economy’s demand for equity-financed technology.

### 4.3 Shock resilience

The model bears another relevant implication: higher efficiency of financial infrastructure helps the economy better absorb exogenous shocks. In the example below, the case is considered of an anticipated shock to output (real shock), but the same result can be shown to hold for nominal shocks as well. Two assumptions need to be added to those previously introduced. First, it is assumed that the conditionality of $\sigma^\rightarrow$ on $w$ in eq. (2) is such that $\sigma^\rightarrow$ increases as $w$ approaches one (Appendix I). Shocks cause suboptimal sales by shortening $\Delta t^0_Q$. Their size increases the amount of $Q$ subject to suboptimal sale. Second, it is assumed that when agents receive signals of increasing future uncertainty (i.e., $w$ tends to one), their probability density function to predict future supply innovations becomes more spread out (Appendix I). Under these assumptions, one can write

$$E[u^\prime(C^\tau_r)] = \phi(C^\tau_r)(\sigma^\rightarrow_t | w_t), \quad \phi^*_c > 0, \phi^*_\sigma > 0$$

thus changing eq. (3) into

$$u^\prime(Q^h_t | w_t, \psi)$$

$$= \sum_{t=1}^\infty \beta^t [1 - d^Q_t(\sigma^\rightarrow | w_t, \psi)] \phi(E(Q^h_t / p^r_t, \sigma^\rightarrow_t | w_t) \prod_p R^Q_t) pr(s^r = s^r_t | w_t) / \sum pr(\cdot)$$

$$= \phi(Q^h_t, p^r_t, \sigma^\rightarrow_t, R^Q_t | w_t, \psi) \quad \phi^*_c < 0, \phi^*_p > 0, \phi^*_\sigma > 0$$
The sign of $\phi_\sigma$ depends on the behavior of both $d^0$ and $pr(\cdot)$ and thus, ultimately, on the liquidity of the asset. Note, however, that $\phi_\sigma (D) > 0$ while $\phi_\sigma (E) < 0$. Positing $w_t = w^0$, rule (8) can now be written as

$$
(8b) \quad \phi(C_t^h, p_t^- \sigma_t^- \mid w^0)(p_t^{C^*_t})^{-1} = \phi(D_t^h, p_t^- \sigma_t^- \mid w^0) = \phi(E_t^h, p_t^- \sigma_t^- \mid w^0, \psi)(p_t^{E^*_t})^{-1} = \mu_t^0
$$

Note that current consumption is conditional on $w^0$ when $\tau > t$. From (8b) and (3b), and recalling that both $(\sigma_t \mid w_t)$ and $pr(s^t = s_{t+1} \mid w_t)$ increase as $w$ approaches one, it follows that an increase in $w_t$ to $w^1 > w^0$ affects equilibrium prices since

$$
(10) \quad \phi(C_t \mid w^1)(p_t^{C^*_t})^{-1} = \phi(D_t \mid w^1) = \mu_t^1 > \mu_t^0 > \phi(E_t \mid w^1, \psi)(p_t^{E^*_t})^{-1}
$$

At date $t$, to re-attain equilibrium (E1)-(E3) and hence (8b), instantaneous prices and inflation must adjust to the new levels $p_t^{C^*_t} > p_t^{C^*_t}$, $r_t^{D^{es}} < r_t^{D^e}$, and $p_t^{E^{es}} < p_t^{E^e}$ and $\pi_t > \pi$ reflecting the agents’ revised expectations over future prices and rates of return as based on new information. With price-elastic supply functions in all three markets, the new equilibrium would be attained at a higher level of current consumption, larger deposit holdings and a smaller stock of equity. Finance to investment would suffer a net reduction.

The result shows that an expected increase in output variability drives risk-averse agents to substitute future with present consumption, and to shift their portfolio towards more liquid assets that better enable them to absorb shocks to consumption with minimum suboptimal asset sales. In fact, nothing ensures that the new equilibrium prices will be attained or sustainable, once attained. Risk perceptions might be such as to lead the agents to deny their money to new supply of less liquid liabilities at whatever price they are offered (rationing).

From the results of the previous section it follows that, ceteris paribus, a more efficient financial infrastructure would require a smaller adjustment in prices and quantities for equilibrium to be re-established after the shock. In other words, what is important to notice is that in an economy with better financial infrastructure the greater efficiency in asset trading - as defined in the context of this study - would better preserve the liquidity of financial assets in the event of expected shocks to output, contain agent runs on quality, and limit the increase in current consumption, as compared to
a system with lower financial efficiency. As a result, rationing phenomena and cuts in investment funds could be limited.

This prediction, for which no empirical evidence is referred to in the next section, will be tested in a subsequent study.

5. Evidence from the literature

Although specific investigation will have to be carried out to provide empirical corroboration of the theoretical predictions of the model above, recent studies seem to bear interesting evidence in their support, thereby encouraging further research along the lines of this study. This section briefly reviews such evidence.

First of all, it is interesting to learn from Demirgüç-Kunt’s and Levine’s (1999) recent cross-sectional analysis of a 60 country data set that the financial systems of countries with less developed financial infrastructure (i.e., less investor-prone law system, poor protection of shareholder and creditor rights, weak contract enforcement, high levels of corruption, poor accounting standards, restrictive banking regulation) tend to be bank-dominated. In particular, the analysis study finds that countries with lower levels of corruption tend to have more market-based financial systems. It also finds that, as financial infrastructure develops, financial systems tend to become more market-based and less bank-dominated.

As regards the relationship between financial development and the quality of capital accumulation, Carlin and Mayer (1998), using data from 27 industries in 20 OECD countries over the 1970-1995 period, find that equity financed industries tend to grow faster, carry out more R&D, and employ higher skilled workers in countries with relatively better accounting standards. Interestingly, under better accounting standards, equity finance industries also tend to undertake less fixed capital accumulation and to invest more in intangibles. This latter result can be explained by the lesser need - made possible by the improved financial infrastructure - for enterprises to accumulate reputational capital in the form of non-salvageable (productive and/or nonproductive) capital assets necessary to signal the enterprises’ commitment to honest and prudent behavior.

Unlike equity-financed industries, Carlin and Mayer find that bank-financed industries grow more slowly in countries with developed financial infrastructure and tend to undertake less R&D. Consistent with Carlin and Mayer’s results, Hoshi et al. (1990, 1990b) also observe that industries dependent on bank finance tend to be physical capital intensive, smoke-stack industries. Furthermore, studying the financing arrangement of 250 US public firms, Houston and James (1995) find that firms with relationships to single banks tend to use less bank debt in
proportion to total debt as their intangible-to-tangible asset ratio (this ratio being proxied by the firms’ market/book value ratio) increases. This contrasts with firms entertaining multiple relationships with banks - that is, where the investor-borrower relationships are not exclusive – in which case the share of bank debt on total debt tends to increase with the relative share of intangibles on total assets.

This could also be interpreted as the result of firms’ strategic choices. In an economy with a well developed financial infrastructure, firms may choose the type of relationship to entertain with investors that best suits their technological requirements. Firms wishing to invest in higher technologies would forego the insurance and advice provided by exclusive relationships in favor of more open and flexible relationships with multiple sources. On the contrary, more traditional investment would induce firms to rely on more exclusive relationships.

Demirgüç-Kunt and Maksimovic (1996) examine the maturity of liabilities in firms from thirty developed and developing countries between 1980 and 1991, and find evidence confirming that firms in developing countries have more limited access to long-term finance due to less developed financial markets and infrastructure. In a subsequent study, they show that an active stock market and a well-developed legal system enhance firms’ access to larger external funds (Demirgüç-Kunt and Maksimovic, 1998). Of particular interest is the highly significant negative correlation between the “law and order” variable used in the study and the rate of return on long-term capital. Although the authors explain this result in term of lower premium on capital due to lower political risk, it also seems perfectly consistent with the transaction cost story of this study and with the study’s theoretical predictions (see section 4.1 above).14

Rajan and Zingales (1998), too, find that industries that are more dependent on external finance tend to grow relatively faster in countries with better financial infrastructure rated according to accounting standards used, types of legal system, and judicial system efficiency. Rajan and Zingales also find that financial infrastructural development has almost twice the economic effect on the growth of the number of establishments in an industry than it has on the growth of their average size. Following the transaction cost approach adopted in this study, such findings may be interpreted by noting that when new firms start out their activity their assets are mainly intangibles. As shown earlier, intangible assets can be financed more easily in economies with more developed financial infrastructure where they trade at lower price discounts.
6. **A crucial question for financial sector reforms in developing economies**

The results of this study may seem to imply that, simply put, banking is bad while non-bank financing is good (or at least better), thus leading to clearcut policy consequences. This is not at all what lies in the author’s mind. To be sure, the model explains why developmental benefits can be attained from improving the economy’s financial infrastructure, even at the expense of lowering the banks’ franchise. One should note, however, that the model assumes as a starting point an emerging economy where banks have *already* achieved a relatively advanced level of development, that is, a point where the economy has already exploited most of the gains from a developed banking sector. In fact, the model could be easily amended to show the benefits from introducing banking in a low-development economy with only commodity-money and small informal lenders.

Also, the benefits from banking are clear when noting, as in section 3.2, that with a poorly developed extrinsic financial infrastructure banks represent the best (if not the only possible) supplement to such missing infrastructure. Indeed, early in development, banks *are* the financial infrastructure of the economy, and internalize the costs and benefits associated with providing financial infrastructural services through exclusive relationships with borrowers and standard debt contacts.

Moreover, the benefits from banking may well go beyond the early stages of development. As Rajan and Zingales (1999) argue, in some cases relationship banking may turn out to be superior to market-based finance. It may help industrial firms with positive franchise to survive temporary distress even if the cost of funding them exceeds their short run debt repayment capacity. Also, relationship banking proves to be more forthcoming than arm’s length competitive financial systems in supporting small and young industrial firms: as relation banks internalize enough of the firms’ franchise, they are more willing than arm’s length financial institutions to make money available to firms, and to subsidize firms intertemporally by lending them cheap when firms are young and by charging them more when they come of age (Petersen and Rajan, 1995). Not to mention the ability of relationship banks to enhance the franchise of their associated firms in economies where capital markets and investors are unsophisticated (De Long, 1991; Chernow, 1997).

Finally, there is no doubt that even in economies with well developed arm’s length competitive financial systems and modern (extrinsic) financial infrastructure, effective banking services are indispensable for the efficient working of finance. The complementarity of banks and non-banks is confirmed by empirical research (Demirgüç-Kunt and Levine, 1996).

What the model above and its results suggest is that “there is life after banking”, and that policymakers must look for *further* considerable gains to be achieved in terms of higher economic
growth and efficiency by replacing in due time the exclusive investor-borrower relationships, typical of traditional banking and early developmental stages, with more arm’s length financial relations and extrinsic financial infrastructure.

In this respect, a serious problem is that, by the time relationship banking produces its full benefits, banks may have accumulated such a strong power and vested interests that likely turn them against any reforms aimed to innovate the economy’s financial relations and infrastructure, as these innovations imply a substantial weakening of the bank rent position. As a result, the reforms could be delayed or blocked until some sweeping financial shocks eventually occur and make them inevitable. The important question - which this study only points to - is that the long-term financial sector development process seems to be characterized by a compelling discontinuity whereby banks need to be supported early in development, only to be - so to say - “weakened” in later stages precisely to foster economic development as higher and better quality growth requires extrinsic financial infrastructure to be introduced at the expense of bank rents. From the policy standpoint, it becomes relevant to determine when and how optimally to generate and manage this discontinuity before it is forced upon the society by some traumatic and costly events such as a banking crisis or the economy’s financial collapse.

Alternatively: Could the discontinuity be avoided altogether by designing financial sector reform programs in ways that would provide for strengthening the banks, while simultaneously building up extrinsic financial infrastructure from the very early stages of economic development? Would this be theoretically consistent and practically viable, considering the importance of relationship banking for successful growth early in development and the banks alleged idiosyncrasy to extrinsic financial infrastructure?

How could such reform programs look like in practice? In very general terms, I can think of two options. First, banks could initially be sustained through appropriate and time-bound financial restraints (Bossone and Promisel, 1998) that would lead them to increase their franchise and accumulate reputational capital. In this case, banks could establish stable relationships with domestic enterprises. The government, however, would have to pre-commit its policy to a strict time deadline after which it would replace restraints with freer competition under strong reputational criteria, and would generate extrinsic financial infrastructure. This option requires, of course, effective specialized supervision and a strong government policy credibility. It is liable to both the risk of encountering (political) resistance from banks and the risk that the government may have an incentive to change policy strategy along the way (dynamic inconsistency).

The second option would be to introduce bank competition and financial infrastructural development simultaneously since the onset of financial sector reform. This option would give an unambiguous signal to the economy as to the government’s intended strategy; it would bear a lesser
risk of dynamic inconsistency in government policy, and would bring efficiency gains more rapidly than under the first option. The downside of this option is that industrial firms would miss the opportunity of benefiting from more stable and durable relationship with banks at a developmental stage where such relationships can be of most valuable use (Cetorelli, 1997).

Important insights to address the “discontinuity” question properly in the case of developing economies can be obtained by assessing the experience of the industrial countries in reforming their financial sector, with particular attention to the response of their banking communities to the reform process.
Appendix I

News, signals, and uncertainty

This approach generalizes the one adopted by Giovannini (1989) to model anticipated shocks. Use $x_i$ to indicate the vector of stochastically-independent real output shocks to the economy. At each date, the agents observe the realization of $x$ and try to anticipate future shocks by using (i.i.d.) current information $w' \in \Omega$. The agents operate a transformation $T$ of $w'$ such that $T: w' \in (\Omega \otimes \mathbb{R}) \rightarrow w \in [0,1]$, which associates to every single bit of information $w'$ a real number ("signal" $w$) in the interval $[0,1]$. In every period the evolution of variable $y$ is governed by the following (conditional) probability distribution function:

$$pr(x_t \mid w_i(t)) = \sum_{n} w_{nt} pr_n(x_t)$$

Where the $w_{nt}$'s are generated by the function $W : w_t \in [0,1] \rightarrow (w_{nt} \in \mathbb{R}^n)$ and satisfy:

1) $0 \leq w_{nt} \leq 1$, $\sum_{t} w_{nt} = 1$
2) $\lim_{w \rightarrow 0} \max(w_{nt}) = 0$ and $\lim_{w \rightarrow 0} \max(w_{nt} ) = 1 \forall t, n \in N$

The $w_{nt}$'s provide a weight structure that is specific to the signal received at each point in time. The greater the uncertainty perceived by the agent, the lower the value of highest weight attributable to the probability of any given shock. The structure of weights associated to every signal by function $W$ determines a probability distribution for each shock $x$. Such distribution is drawn from a set of distribution functions $pr_n(\cdot)$'s obeying the following restrictions:

i) $E(x_t pr(x_t \mid w_i(t))) = E(x_t)$, $\forall w$, that is, all distributions are mean-preserving;
ii) $pr_j(x_t \mid w_i(t)) - pr_j(x_t \mid w_j(t)) = MPS(x_t)$, $\forall k, j$, that is, the distribution spreads are mean-preserving (Rothschild and Stiglitz, 1970)
iii) $\lim_{w \rightarrow 0} [pr(x_t \mid w_i(t)) - pr(E(x) \mid w_i)] = 0$, that is, the probability density function of any given shock $x_i$ becomes more spread out as the signal approaches one.

The rationale for this formal structure is that for any given signal received, the agent forms a specific conjecture as to the possible occurrence of a future supply innovations. The structure of weights assumed represents the degree of belief (or confidence) that the agent attaches to such conjectures. The value of each received signal $w$ reflects the degree of uncertainty perceived by the agents: as $w$ increases, the weights change so as to make any conjecture on future shocks weaker and, thus, any prediction more tenuous.
Appendix II

Asset realization and asset price discount

As the date of asset realization falls closer to planning date \( t \), the risk of suboptimal sale increases. There is thus a link between the probability of date-event \( s'_t \), the proximity of \( \tau \) to \( t \), and the size of \( d^Q \). To show this, consider two extreme cases by solving equation (2) for some critical values of \( d^Q \) and \( pr(\cdot) \), assuming \( Q \) has maturity \( T \), \( d^Q^* = 0 \), \( r^Q = \text{constant} \), and \( \pi = 0 \).

Case 1): if at time 0: \( pr(s^c = s^c_T) = 0, \tau = 0, T - 1 \), that is, no shock to consumption is anticipated during the life of the asset (which is equivalent to the case where \( pr(s^c = s^c_T) = 1 \)), and the agent is certain that she will not have to liquidate \( Q \) at a discount (\( d^Q = 0 \)), the current marginal utility of \( Q \) is

\[
\nu = \beta^T_0 E[u^*(P^Q_0 Q_0 / p^*_T)R^{0T}]
\]

Case 2): If at time 0: \( pr((s^c = s^c_T), \tau / \Delta t^*_Q \rightarrow 0] = 1 \), that is, a shock to consumption is anticipated for date \( \tau \), and \( \tau \) is such that the agent will have to sell the asset suboptimally, then \( d^Q \rightarrow 1 \) and \( \nu \rightarrow 0 \).

These examples represent benchmarks for more realistic cases. For, at times of higher output variability, asset price discounts are likely to increase to the extent that the subjective probability of having to realize illiquid assets at a short notice at each date is higher, and agents find themselves compelled to sell the assets suboptimally. Under such circumstances, the current marginal utility of the assets involved tends to decrease.
References


-------------------------------------


Endnotes

1 Fry (1995), Ch. 1, and OECD (1997).
4 Transaction costs here thus refer to the costs incurred by the agents to: i) search for trustworthy counterparts to trade; ii) assess the quality of the assets and commodities submitted to trade; iii) ascertain the trustworthiness of trading counterparts; iv) fix all legal and property-right related issues; and v) monitor and enforce contracts.
5 In fact, where banks have a strong quasi-monopolistic power, they can extract rents from transactors by charging non-competitive fees on payment-related transactions. Yet, at least part of such rents reflect a premium that transactors are willing to pay in return for safety on transactions and lower costs for assessing payors’ trustworthiness and quality of assets.
6 The rationale for asset price discounting under incomplete trust is discussed in Bossone (1999) cited in footnote 3. For readers’ convenience, the mechanics of discounting is described in this section.
7 The quality of an asset reflects the asset’s liquidity and store-of-value capacity. To assess the asset’s quality, the agents need to gather information on: i) the adherence of the asset’s ask-price to its fully informational level, which in turn reflects the asset’s relative scarcity (in case of a commodity), or its marginal efficiency (in case of a capital good); ii) the adherence of the asset’s characteristics to those claimed by the seller; iii) the conditions of the market where the asset is traded. Underlying the possibility of non-adherence under items i) and ii) is the existence of asymmetric information and of limited trust between traders.
8 This concept draws on, and integrates, Lippman and McCall’s (1986) concept of optimal search time. Optimal transaction time period is determined by the trading context as defined by the institutional, legal and
technological settings for trading, and as shaped by the trading customs and practices prevailing in the society considered.

9 See also this argument in Chernow’s (1997) explanation of the fall of investment banking in the US after the Securities Act of 1933.

10 Note that date-events $s_t^c$ are mutually independent across $t$ so that $\sum_t pr(s_t^c = s_t^c) = 1$ does not necessarily hold. At an extreme, for instance, one could have that $pr_t(\cdot) = 1$ for each and all $t$’s.

11 The two cases would be reflected in the following two disequilibrium relations. For the first case:

$$V(C_t^{hs})(p_t^{cs})^{-1} = \mu_t^1 < \mu_t < V(D_t^{hs}, R_t^{E-}) < V(E_t^{hs}, R_t^{E-} | \psi^1)(p_t^{E*})^{-1}$$

and for the second case:

$$V(C_t^{hs})(p_t^{cs})^{-1} = \mu_t^1 < \mu_t < V(D_t^{hs}, R_t^{E-}) = V(E_t^{hs}, R_t^{E-} | \psi^1)(p_t^{E*})^{-1}$$

12 F. Zahir and I are planning to carry out such an investigation on a cross-country basis. Results should be available by year-end.

13 See Bossone (1999), cited in footnote 3.

14 Contrary to our model, Demirgüç-Kunt and Maksimovic (1998) find no correlation between law and order and the proportion of firm asset growth finance with external equity, while they do find that law and order are highly positively correlated both with the size of the banking sector and the proportion of firm asset growth financed with long-term debt.

15 See Bossone (1999), cited in footnote 3.