

Chapter 6
Growth and Finance

October 19, 2006

1 Introduction

Financial markets and financial intermediaries are important for economic growth, because in various ways they facilitate the investments in capital and technology that underlie long-term growth. For one thing, people are willing to save more, and hence make more available to investors, in a country with efficient and trustworthy banks than in a country where banks are likely to waste their depositors’s wealth through bad loans or even swindles. Banks can also help by risk-pooling. That is, by collecting savings from many people and investing them in a large diversified range of projects a bank allows even small savers to take advantage of the law of large numbers and get a reasonably safe rate of return; the losses on bad projects will tend to be offset by the gains on good projects. In addition, well functioning banks can channel savings towards the most efficient uses. That is, since the people with the best investment projects are not necessarily the ones with the most savings, intermediaries are needed to identify who has the best projects and to match them up with the surplus funds of savers. Finally, and perhaps most importantly, banks can also help to alleviate agency problems by monitoring investors and making sure that they are making productive use of their loans rather than spending them on private consumption or otherwise defrauding the ultimate lenders. Many of these roles are also served by stock markets, private equity firms and venture capitalists, all of which help to identify, finance and monitor good investment projects.

In his excellent survey article on Finance and Growth for the forthcoming Handbook of Economic Growth, Ross Levine summarizes as follows the existing research on this topic: “Taken as a whole, the bulk of existing research suggests that (1) countries with better functioning banks and markets grow faster; (2) simultaneity bias does not seem to drive these conclusions, and (3) better functioning financial systems ease the external financing constraints that impede firm and industrial expansion, suggesting that this is one mechanism
through which financial development matters for growth”. In this chapter, we introduce financial constraints into the Schumpeterian growth framework, and show explicitly how growth is enhanced by intermediaries that provide external finance to innovators by channeling savings and mitigating agency problems. We provide two alternative models for this; in the first one banks provide screening services to identify good projects, and in the second one banks provide ex post monitoring to make it difficult for the borrower to abscond with borrowed money. Then we describe some of the voluminous empirical evidence underlying Levine’s summary statement of the role of finance in the growth process.

2 Innovation and growth with financial constraints

In this section, we introduce credit constraints in the multi-sector Schumpeterian growth model of the previous chapter.

2.1 Basic setup

The economy has a fixed population $L$, which we normalize to unity. Everyone lives for two periods, being endowed with one unit of labor services in the first period and none in the second, with a utility function linear in consumption: $U = c_1 + \frac{1}{1+r}c_2$. The rate of time preference $r$ will be the equilibrium rate of interest.

There is one final good, produced by labor and a continuum of intermediate inputs according to:

$$Y_t = L^{1-\alpha} \int_0^1 A_{it}^{1-\alpha} x_{it}^\alpha di, \quad 0 < \alpha < 1$$

where $x_{it}$ is the input of the latest version of intermediate good $i$ and $A_{it}$ is the productivity parameter associated with it. The final good is used for consumption, as an input to R&D, and also as an input to the production of intermediate goods.

Each intermediate good $i$ is produced by an individual born each period $t - 1$ and who can potentially innovate at time $t$. Let $\mu$ be the probability that innovation in any sector $i$ succeeds at time $t$. (In equilibrium this probability will be constant over time and the same for all sectors $i$). An innovator starts with the average productivity of period $t - 1$:

$$A_{t-1} = \int_0^1 A_{i,t-1} di$$

and a successful innovation allows the sector to improve on this by the factor $\gamma > 1$. So in every period $t$ the fraction $\mu$ of sectors (those who have innovated) will have productivity
γA_{t-1} while the remaining fraction \((1 - μ)\) have \(A_{t-1}\). The average across all sectors is therefore:

\[ A_t = μγA_{t-1} + (1 - μ)A_{t-1}, \]

implying that the growth rate of average productivity is:

\[ g = \frac{A_t - A_{t-1}}{A_{t-1}} = μ(γ - 1). \] (2)

In each intermediate sector where an innovation has just occurred, the incumbent is able to produce any amount of the intermediate good one for one with the final good as input. In addition, in every intermediate sector there is an unlimited number of firms, the competitive fringe, that can produce the same quality of that intermediate good at a unit cost of \(χ\) units of the final good, with \(χ \in (1, α^{-1})\).

The final good, which we take as the numeraire, is produced under perfect competition, so the price of each intermediate good equals its marginal product:

\[ p_{it} = αA_{it}^{1-α}x_{it}^{α-1}. \]

In sectors where an innovation has just occurred, the incumbent will be the only producer, at a price equal to the unit cost of the competitive fringe, whereas in non-innovating sectors where the most recent incumbent is dead, production will take place under perfect competition with a price equal to the unit cost of each producer. In either event the price will be

\[ p_{it} = χ, \]

so from the last two equations the equilibrium quantity of each intermediate product will be:

\[ x_{it} = (α/χ)^{1-α} A_{it} \]

Thus an unsuccessful innovator will earn zero profits next period, whereas the profit of an incumbent will be

\[ π_{it} = δA_{t-1}, \]

where \(δ = γ(χ - 1)(α/χ)^{1-α} \).

From the above, gross output of the final good in equilibrium will be:

\[ Y_t = θA_t \] (3)

where \(θ = (α/χ)^{1-α}\) and \(A_t\) is again the average productivity parameter across all interme-
2.2 Innovation technology and growth without credit constraint

Suppose that the R&D cost of innovating with probability $\mu$ is equal to:

$$N_t = A_{t-1} \phi (\mu^2 / 2), \quad \phi > 0$$

(4)

where this cost is incurred in units of the final good. The multiplicative term $A_{t-1}$ reflects the fact that the further ahead the current technology is, the more costly it is to innovate upon it.

This means that by spending $N_t$ units of the final good in R&D at date $t$, a firm innovates with probability

$$\mu_t(N_t) = \sqrt{\frac{2N_t}{A_{t-1}\phi}}.$$  

(5)

In equilibrium $\mu$ will be chosen so as to maximize the expected net payoff:

$$\mu \delta A_{t-1} - A_{t-1} \phi \mu^2 / 2,$$

(6)

hence the equilibrium probability of innovation is:

$$\mu^* = \delta / \phi.$$

From this and (2), the corresponding equilibrium growth rate is:

$$g = \mu^*(\gamma - 1) = (\delta / \phi)(\gamma - 1).$$

2.3 Credit constraints: A model with ex ante screening

Each innovator at date $t$ is a young person with access to the wage income $w_{t-1}$. Thus to invest $N_t$ in an R&D project she must go to a bank and borrow $L = N_t - w_{t-1}$, which we suppose is strictly positive. Suppose (following King and Levine, 1993) that there is also a number of other people seeking to finance projects but that these projects are in fact not feasible. Let $\varphi$ be the probability that a borrower coming to a bank is capable (has a feasible project) while $1 - \varphi$ is the probability that the borrower’s project will yield no payoff at all.

A bank can determine whether or not a give project that involves spending $N_t$ on R&D is feasible by paying a cost equal to $f N_t$ in units of the final good. Then it will require a repayment of $f N_t / \varphi$ from each feasible project in order to break even; the combined payoff
to the (capable) innovator and her bank will be the expected profit of a successful innovation minus the R&D cost minus the screening cost:

$$\mu\delta A_{t-1} - N_t - f N_t/\phi$$

So if the R&D cost of innovating with probability $\mu$ is again given by equation (4) above, then the equilibrium value of $\mu$ will be chosen to maximize:

$$\mu\delta A_{t-1} - (1 + f/\phi) A_{t-1}\phi \mu^2/2$$

which results in an equilibrium probability of innovation equal to:

$$\mu^* = \frac{\delta}{(1 + f/\phi) \phi}.$$  

From this and (2), the corresponding equilibrium growth rate is:

$$g = \frac{\delta}{(1 + f/\phi) \phi} (\gamma - 1).$$

Therefore the higher the agency cost $f$, the lower will be the frequency of innovations and the lower will be the equilibrium growth rate.

### 2.4 A model with ex post monitoring and moral hazard

Our second model of credit constraints and growth focuses on ex post monitoring following Aghion, Banerjee and Piketty (1999) and their emphasis on the notion of credit multiplier.

#### 2.4.1 Credit multiplier and R&D investment

Again, each innovator is a young person with access to the wage income $w_{t-1}$, who must borrow $L = N_t - w_{t-1}$. Now, suppose that what makes it difficult to borrow is that at a cost $c N_t$ the innovator can hide the result of a successful innovation and thereby avoid repaying her creditors, where $0 < c < 1$. This cost as an indicator of the degree of creditor protection. In countries where laws and institutions make fraud a costly option creditors are better protected and therefore, as we shall see, credit is more readily available to entrepreneurs.

The innovator must pay the hiding cost at the beginning of the period, when she decides whether or not to be dishonest. She will do so when it is in her self interest, namely when
the following incentive-compatibility constraint is violated:

\[ cN_t \geq \mu_t(N_t)R(N_t - w_{t-1}) \]  \hspace{1cm} (7)

where \( R \) is the interest factor on the loan and \( \mu_t(N_t) \) is the probability of innovating at date \( t \) given the R&D investment \( N_t \), as determined by equation (5) above. The right hand side of (7) is the expected saving from deciding to be dishonest when investing at the rate \( N_t \).

The only potential lenders in this overlapping generation model are other young people, who will lend only if offered an expected rate of return equal to \( r \). Thus the interest factor on the loan in equilibrium must satisfy not only the incentive-compatibility condition (7) but also the arbitrage condition:

\[ \mu_t(N_t)R = 1 + r \]

so that the incentive-compatibility condition boils down to an upper limit on the entrepreneur’s investment:

\[ N_t \leq \frac{1 + r}{1 + r - c}w_{t-1} = \nu w_{t-1}. \]

The parameter \( \nu \) is commonly referred to as the credit multiplier. One can immediately see that a higher cost \( c \) of hiding ex post innovation revenue implies a higher value of \( \nu \).

Moreover, one can easily show that \( \nu \) or \( c \) map one to one with the ratio of private credit to GDP used to measure financial development in the empirical studies surveyed in the next section. Indeed, the ratio of intermediary lending to GDP is equal to:

\[ F = \frac{N - w}{Y} \]

where \( N \) is R&D expenditure, so that \( N - w \) is the amount of R&D expenditure that is financed through intermediaries. In a country where financial constraints are binding:

\[ F = (\nu - 1) \frac{w}{Y}. \]

Given that wages constitute about 70 percent of per-capita GDP in most countries, this implies that:

\[ F \simeq 0.7\nu \]

which is strictly proportional to the credit multiplier \( \nu \). Thus the parameter \( \nu \) truly captures the economy’s degree of financial development.
2.4.2 Innovation and growth under binding credit constraint

This limit will be binding if the unconstrained optimal investment $\mu^*$ violates it, that is, whenever:

$$\mu_t(N_t) = \sqrt{\frac{2\nu w_{t-1}}{A_{t-1}\phi}} < \mu^* = \frac{\delta}{\phi}. \tag{8}$$

Using the fact that under the Cobb-Douglas specification in (1), the equilibrium wage, $w_{t-1}$, is equal to $(1 - \alpha)$ times final output $Y_{t-1}$ at date $t - 1$, and that final output is itself proportional to aggregate productivity $A_{t-1}$ in equilibrium at date $t - 1$ by equation (3), we have:

$$w_{t-1} = \omega A_{t-1},$$

with

$$\omega = (1 - \alpha)\theta.$$

But then we can simply reexpress the innovation constraint (8) as:

$$\nu < \frac{\delta^2}{2\phi\omega}. \tag{9}$$

In particular, we see that the credit constraint is more likely to be binding when: (i) financial development, as measured by $\nu$, is low; (ii) the marginal cost of R&D, as captured by $\phi$, is low, so that the unconstrained equilibrium R&D investment is high; (iii) entrepreneurs’ initial wealth as a fraction of aggregate output, namely $\omega$, is low.

Whenever (9) holds, the equilibrium growth rate is given by:

$$g^e = \sqrt{\frac{2\nu \omega}{\phi}}(\gamma - 1),$$

which is monotonically increasing in financial development as measured by $\nu$.

3 The empirical findings: Levine’s survey in a nutshell

Most of the empirical literature on finance and growth, has been concerned with cross-country or panel regressions of the form

$$g_i = \alpha_0 + \alpha_1 F_i + \alpha_2 X_i + \varepsilon_i,$$

where $g_i$ is the average growth rate in country $i$ during the period or subperiod, $F_i$ is the country’s level of financial development (either at the beginning of the period, or averaged
over the period), $X_i$ is a vector of controls (policy variables, education, political stability, initial income per capita, etc) and $\varepsilon_i$ is a noise term.

As well explained in Levine’s survey, empirical papers on finance and growth differ in terms of: (i) whether they look at cross country data (like King and Levine (1993) and subsequent work by Levine and coauthors) or at cross industry data like Rajan and Zingales (1998) or at cross regional data like Guiso, Sapienza and Zingales (2002) or at firm level data like Demirgüç-Kunt and Maksimovic (1998); (ii) how is $F_i$ measured: by the ratio of bank credit to GDP, or by indicators of stock market development, or if it is also interacted like in Rajan and Zingales (1998) with a measure of external financial dependence of the industry; (iii) whether one looks at cross section or at panel data; (iv) whether or not one instruments for financial development.

### 3.1 Cross-country

King and Levine (1993)\(^1\) consider a broader sample of 77 countries over the period 1960 to 1989. They regress average growth of per capita GDP or average growth in TFP over financial development and a number of control variables, as specified in the above equation. The controls include: initial income per capita, education measures, indexes of political stability, and policy indicators. Financial development is measured in three possible ways: (i) the ratio between the liquid liabilities of the financial system and GDP; (ii) the ratio of commercial bank credit to bank credit plus central bank’s domestic assets (this measure performs generally more poorly than the others); (iii) the ratio of credit to private enterprises to GDP. Each of these measures is averaged over the period 1960-1989. The cross-country regressions in Table 1 shows a large and significant correlation between productivity growth and the above measures of financial development measured as specified above.

\begin{center}
\textit{TABLE 1 HERE}
\end{center}

To make sure they capture the causal relationship from finance to growth and not the reverse, King and Levine repeat the same regression exercise but using initial 1960 values of financial depth instead of their average over the whole period. This regression also shows

\(^1\)Levine (2005) attributes the first empirical analysis on finance and growth to Goldsmith (1969). Goldsmith uses cross-country data over the period 1860-1963 to regress average growth on financial development as measured by the size of the financial intermediary sector (measured by the value of its assets) over GDP, and finds a positive correlation between financial development and output growth. As well explained by Levine, this study has its limits: no controls in the regression, no instrumentation to address potential causality issues, the left hand side variable is output growth instead of productivity growth, and the sample consists of 35 countries only. It is these limitations that King and Levine address in their seminal work in 1993.
a positive and significant correlation between financial development and growth, which now suggests that "financial development in 1960 is a good predictor of economic growth over the next 30 years".

Subsequently, Levine and Zervos (1998) concentrate on the nature of financial sectors, and in particular the importance of stock market development and stock market "liquidity". In particular, Levine and Zervos consider what they call the "turnover ratio", namely the total value of currently traded shares over the total value of listed shares, and based on a cross-country regression involving 42 countries over the period 1976-1993, they find that both, the initial level of bank credit and the initial level of this turnover ratio in 1976, show a positive and significant correlation with average productivity growth over the period 1976-1993, as we see in Table 2 below.

TABLE 2 HERE

One may object to the measures of financial development used by Levine and his co-authors, however this is the best that can be done while remaining at cross-country level.

A more serious objection is causality: what tells us that these positive correlations do not reflect, either the fact that financial development occurs in prediction of forthcoming growth, or the fact that a third variable, call it institutional development (e.g, measured by property rights protection), causes both, higher growth and higher financial development. To address this endogeneity problem, Levine (1998, 1999) and Levine, Loayza and Beck (2000) use the legal origins indicators of La Porta et al (1998) as instruments for financial development. Thus the regression exercise now involves a first stage where financial development is regressed over dummy variables for Anglo-Saxon, French and German legal origins (against Scandinavian legal origins) respectively, and a second stage regression where average productivity growth is regressed over predicted financial development as derived from the first stage regression and the same control variables as before. In particular, Levine et al (2000) obtain a strongly positive and significant correlation between predicted financial development and average productivity growth over the period 1960-1995, as shown in Table 3.

TABLE 3 HERE

Levine et al (2000) go even further by performing panel cross-country regressions in which the period 1960-1995 is subdivided in five year subperiods, and where, for each five-year subperiod, average productivity growth over the subperiod is regressed over current and past financial development, controlling for country fixed effects. And again, they find
positive and significant correlations between (current and lagged) financial development and average productivity growth during the subperiod.

Because they move from cross-country to cross-regional analysis within a country (Italy), Guiso, Sapienza and Zingales (2002) can construct more precise measures of financial development and show that financial development as they measure it is an important determinant of cross-regional convergence. More specifically, GSZ construct their measure of regional financial development by estimating a linear probability model in which they regress the probability that individuals be denied access to credit (they obtain information about individual access to credit from the Survey of Households Income and Wealth, which also provides information on the region to which each individual belongs) over regional dummies and a set of control variables. The coefficients on the regional dummies are the measures of regional financial development, which GSZ instrument using the regional composition of banking branches in 1936.2

3.2 Cross-industry

The pioneering attempt at getting at a more microeconomic level by looking at cross-industry comparisons across countries, is by Rajan and Zingales (1998). Their insight is that growth in industries that rely more heavily on external finance, should benefit more from higher financial development than growth in industries that do not rely so much on external finance. The problem is to identify those industries that are more prone to rely on external finance than other industries.

Rajan and Zingales regress average growth of value added of industry $k$ in country $i$ over: (i) country and industry dummies; (ii) the share of industry $k$ in total manufacturing in country $i$; (iii) the interaction between financial development (measured by stock market capitalization plus domestic credit over GDP) in country $i$ and industry $k$’s dependence upon external finance (measured by the fraction of capital costs not financed internally in that same industry in the US). The underlying idea is that firms are not financially constrained in the US, so that this measure of external dependence can be thought of as being independent from financial development and to depend instead upon technological factors only. Rajan and Zingales do not include financial development independently, as this would create collinearity with the country dummies.

Using a sample of 36 industries in 42 countries, Rajan and Zingales find an interaction coefficient between external dependence and their measure of financial development, which is positive and highly significant at the one percent level), thereby providing strong evidence 

2 The year 1936 corresponds to the enactment of a law restricting subsequent entry into the banking sector.
to the effect that higher financial development enhances growth in those industries that rely more heavily on external finance. Table 4 summarizes their main findings.

**TABLE 4 HERE**

Building upon the Rajan-Zingales methodology, Beck, Demirgüç-Kunt, Laeven and Levine (2004) use cross-country/cross-industry data to look at the effect on productivity growth of the interaction between financial development and the average size of firms in the corresponding industry in the US (again relying on the implicit assumption that only technological factors, not financial market frictions, determine this average size in the US). They find that higher financial development enhances growth in those industries that comprise a higher fraction of small firms. This result is consistent with previous work by Bernanke, Gertler and Gilchrist (1989) suggesting that smaller firms face tighter credit constraints than large firms.

### 4 Conclusion

In this chapter, we have introduced financial constraints into the Schumpeterian growth framework, and confronted the prediction that growth should increase with financial development with existing empirical evidence from cross-country and cross-industry data.

More recent work has looked at cross-firm data. In particular, Demirgüç-Kunt and Maksimovic (1998), henceforth DM, have analyzed the extent to which long-term debt and outside equity financing can foster firm growth. To this end, they first computed the growth rate of firms that would not have access to long-term debt or outside equity (that is, the growth rate of firms that only rely on retained earnings and short-term debt); then they calculated the fraction of firms that grow faster than the no-outside-finance rate; this they interpret as being the fraction of firms that rely on outside finance, and DM compute this fraction \( f_i \) for each country.

Then, using a sample comprising all large publicly traded manufacturing firms in each of 26 countries, DM regress the fraction \( f_i \) of firms that grow faster than the no-outside-finance rate, over financial development (measured either by the ratio of market capitalization to GDP, or by the turnover ratio of Levine and Zervos to capture the liquidity of stock market, or by the ratio of bank assets to GDP to capture the size of the banking sector) and control variables. The main finding in DM is that the turnover ratio and the bank assets to GDP ratio are both positively and significantly correlated with \( f_i \).

Another important piece of work on credit constraints at firm level, is the innovative paper by Banerjee and Duflo (2004), which uses firms’ investment response to an exogenous
policy change affecting the amount of subsidized directed credit, to assess the importance of credit constraints faced by firms. The underlying idea is that an unconstrained firm would respond to such a policy change by simply substituting directed credit for (unsubsidized) market credit, but without changing capital investment (which in that case would only be determined by rate of return considerations). The policy change is that the limit on total capital investment for a firm to belong to the so-called priority sector eligible for subsidized credit, was raised substantially in 1998 and then lowered back in 2000. Banerjee and Duflo then show that bank lending and firm revenues increased for the newly targeted firms immediately after 1998, and then decreased in the years after the 2000 policy change, thereby providing evidence to the effect that those firms were indeed credit constrained.

In the next chapter we will go one step further and show that by interacting a country’s level of financial development with its level of technological development, one can account for observed convergence and divergence patterns. In chapter X, we will again bring financial constraints into the picture, but this time to analyze the interplay between macroeconomic volatility and growth, and the potentially growth-enhancing role of countercyclical fiscal policies.