SHORT-RUN COSTS OF FINANCIAL MARKET DEVELOPMENT IN INDUSTRIALIZED ECONOMIES

Thomas Lindh
Uppsala University and *Institute of Futures Study*

INTRODUCTION

The financial sector share in most industrialized economies has more or less doubled in the last few decades. The expansion is associated with decreasing, rather than increasing, growth rates over most of this period. Only in the last decade do we see a trend towards increased productivity growth. This is puzzling, since there is a well-documented positive correlation between financial and economic development, [Goldsmith, 1958; 1969]. Growth-enhancing effects of financial markets seem to show up only in the long run while the short-run effect is to temporarily inhibit growth.

A tentative explanation is offered here. An extension of the financial sector requires organization of new markets. This can only be done in discrete steps due to indivisibilities. A fixed cost associated with these indivisibilities temporarily raises the relative cost of financial services. Thus, resources are diverted from production causing a temporary fall in the growth rate. However, increased risk sharing favors specialization, and as the economy adapts growth recovers. In the growing economy the benefits of pooling risks sooner or later again outweigh the fixed cost of a further extension of the financial markets, and so forth.

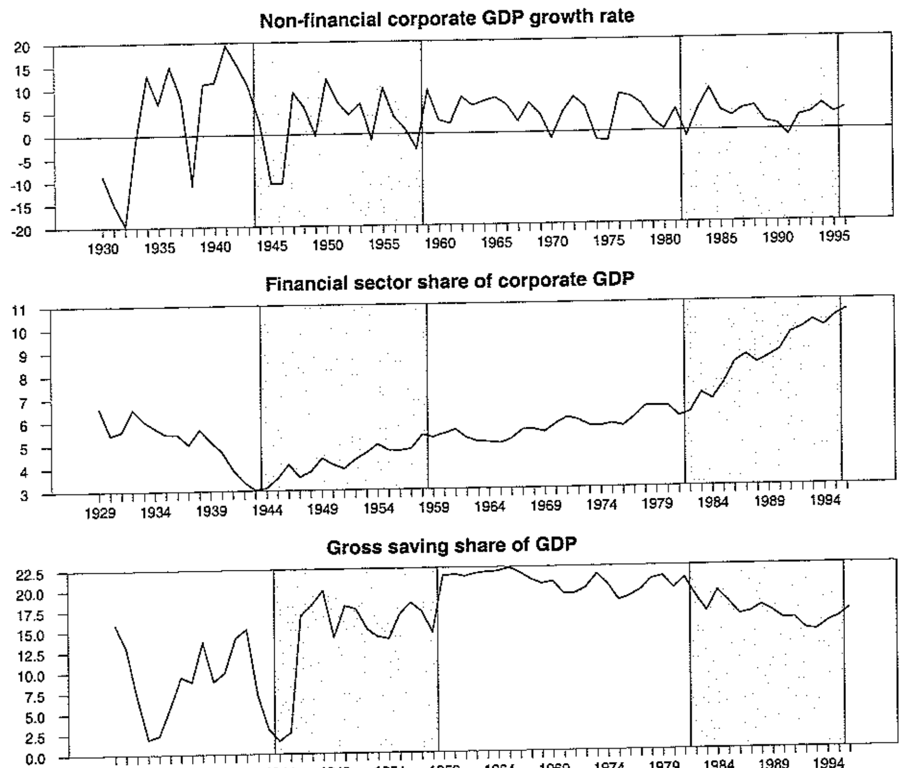
The model developed in this paper reconciles long-run growth-enhancing financial development with short-run slow-down effects. Since the mechanisms are embedded in an endogenous growth model, saving rates take on a much more important role than in the traditional Solow growth model, the latter only permits transitory effects on growth while the former allows for permanent effects. Saving is negatively affected by increasing costs in the transformation of saving to investment, and probably also by decreases in precautionary saving. This is consistent with another stylized fact: the postwar decrease in national saving rates.

This paper provides a consistent theoretical framework to explain the dynamics of a staggered development with regime shifts in the relation between the financial sector and the growth of output in industrialized countries. The paper does not attempt to empirically test this hypothesis. There is, however, empirical evidence of regime shifts in the relation between the financial sector on one hand, and saving and growth, on the other hand (for the United States see Jacobson et al. [1998] and for the G-7 countries Lindh and Lindström [1997]). The next section reviews and summarizes the empirical and theoretical background. The conjectured explanation is formally modeled. The final section concludes and summarizes.

Thomas Lindh: Department of Economics, Uppsala University, Box 513, SE-751 20, Uppsala, Sweden. E-mail: thomas.lindh@nek.uu.se

Eastern Economic Journal, Vol. 26, No. 2, Spring 2000

FIGURE 1
United States Annual Data 1929-1996



Shaded areas mark expansionary phases for the financial sector share. Source: National Income and Product Accounts, U.S. Department of Commerce [1992; 1997].

FINANCIAL DEVELOPMENT, GROWTH AND SAVING

The first subsection briefly illustrates some empirical facts in U.S. data. Theoretical and empirical research that is relevant for the paper is surveyed in the following two subsections.

An illustration: United States 1929-1996

The middle panel in Figure 1 shows the ratio of financial corporate GDP to total corporate GDP in the United States. The growth rates of non-financial corporate GDP are plotted in the upper panel and the gross saving share of GDP in the lower panel.[1]

The financial sector share shows quite different trend behavior in different periods. One can discern a division of roughly four periods between 1929 and 1991. The first, 1929-1943, was a period of contraction, including the Great Depression and ending at the outbreak of World War II. The financial sector share shrank to less than half of its 1929 value in 1943. Up to around 1958 the financial share rapidly expanded

TABLE 1
Sample Means and Estimated Linear Trends in Sub-periods

	Contraction 1929-1943	Expansion I 1944-1958	Tranquility 1959-1981	Expansion II 1982-1996
Financial Share Trend	−0.20	0.12	0.06	0.30
	(0.03)	(0.02)	(0.01)	(0.02)
Mean Growth Rate	4.0	2.4	4.4	3.4
	(3.3)	(1.2)	(0.7)	(0.7)
Growth Trend	2.1	0.3	−0.14	−0.02
	(0.6)	(0.4)	(0.10)	(0.16)
Mean Saving Share	9.1	14.6	20.7	16.5
	(1.2)	(1.4)	(0.2)	(0.4)
Saving Trend	−0.01	0.6	−0.1	−0.2
	(0.3)	(0.3)	(0.03)	(0.06)

Standard errors of the estimates are reported in parentheses. All measured variables are in percentages.

again, recovering to approximately its 1930 value. Then followed a period of more tranquil growth in the share 1959-1981. In the last fifteen years the financial share again expanded rapidly (from around 7 to 10.5 percent).

The correlation coefficient between the financial sector share and the growth rate over the whole period is −0.14. Although this overall estimate is statistically insignificant, the correlation coefficients are negative in each of the sub-periods marked in Figure 1. In the contraction phase 1929-1943 the correlation is −0.63 and is statistically significant at the 5 percent level. In the more tranquil phase 1959-1981 the correlation is −0.35 which is borderline significant. In the expansionary phases of the financial sector share 1944-1958 and 1982-1996 the correlation is still negative but close to zero and statistically insignificant.

A correlation is no proof of a relation, but it is still surprising in view of standard results showing a positive relation between financial development and growth. How come? The value added in the financial sector would seem intuitively to be related to financial development. High average growth rates in the 1960s and 1990s *follow* the expansions of the financial share in the 1950s and 1980s with a quite considerable lag. The gross saving share in the lower panel also seems to go down initially as the financial sector expands.

Descriptive statistics in Table 1 show that average growth rates in expansionary phases are lower than in the preceding periods. The linear growth trend estimate is significantly positive only in the contraction phase while the hypothesis of a constant growth rate cannot be rejected for the other periods. The corresponding saving pattern is very volatile up to the 1950s and the first expansionary phase, but does show a tendency to decrease in the second financial expansion period. From these data it would thus be reasonable to conjecture that financial sector expansions are associated with lower growth and possibly also a decrease in saving rates.[2]

Theoretical background

Financial development can influence growth in three distinct ways: by raising the proportion of saving actually invested; by raising marginal productivity in the economy; and by influencing the private saving rate. The first mechanism depends on the efficiency of financial intermediation, (i.e., the fraction of saving absorbed to pay for financial intermediation services). The second mechanism works by improving the allocation of capital through information pooling [Greenwood and Jovanovic, 1990] or risk pooling, (e.g. liquidity risks [Bencivenga and Smith, 1991]). The third mechanism has ambiguous effects on growth since saving may go either way. If consumers have utility functions with a positive third derivative, precautionary saving decreases. When, as will be assumed in the model of this paper, the utility function is of the constant relative risk aversion (CRRA) type, the relative risk aversion parameter is then greater than one. Financial markets may also ease liquidity constraints by providing consumption credit. On the other hand, financial development may also increase the rate of return and therefore boost saving. Devereux and Smith [1994] and Obstfeld [1994] study models incorporating these mechanisms. Pagano [1993] provides a brief overview and summary of the literature about financial development and growth. Levine [1997] contains a more extensive survey.

If there are inherent fixed costs in establishing financial markets—as is very likely since it entails new information networks—extensions will take place discretely. Saint-Paul [1992a] models how fixed financial market costs that imply a trade off between risky specialization and financial costs for diversifying income risks may induce a poverty trap in an endogenous growth context. The model developed in the next section takes Saint-Paul's model as its starting point. Alternatively, indivisibilities in the production technology may require a threshold level of financial development in order to be insurable [Acemoglu and Zilibotti, 1997]. Yet another source for scale effects caused by informational asymmetries is explored by Khan [1999].

Different theoretical approaches tend—as Pagano [1993] points out—to yield similar implications for the relations between finance, saving and growth on the macro level. Growth is enhanced by financial development through a variety of mechanisms that improve productive efficiency or increase the volume of savings. Mechanisms, such as screening and monitoring, higher transactional efficiency etc. working through liquidity risks, informational pooling, and so forth have been proposed in the literature. Growth is negatively affected by the resource cost of financial markets and if risk aversion is high enough, possibly also by a negative effect on precautionary savings. Moreover, growth tends to feed back into increased demand for financial insurance by increasing aggregate risks.

While most of the literature concerns a long-run steady state, what we observe is a variation in the relations over different periods. Saint-Paul [1992a] and Acemoglu and Zilibotti [1997] allow for a one-shot take-off relevant for the discussion of how developing countries may get out of poverty traps by tapping into the international financial markets. This paper extends this work to a sequential process of market extensions, which makes the same ideas relevant for developed countries. As illustrated in the U.S. time series there are clearly distinct periods of financial development. The relation between growth and saving on the one hand, and the financial sector on the other, differ considerably in periods of financial expansion from periods of calmer development or contraction.

Empirical Evidence

King and Levine [1993a and 1993b] examine cross-country evidence for 77 countries from 1960-1989 using four different measures of financial development and find a strong positive correlation with average growth rates. King and Levine [1993b] use predetermined levels of financial variables to conclude that the correlation is a long-run positive growth effect from financial development. Fernandez and Galetovic [1994] replicate the King and Levine study, but split the sample into OECD countries and non-OECD countries, finding statistically insignificant coefficients for the OECD sample, indicating that for mature economies growth-enhancing effects are less.[3] The OECD countries, of course, also have more developed financial markets and the cross-section averages hide a considerable expansion over the time period in question.

The causality of the empirical positive long-run relation between growth and financial development is still a matter of controversy. It is also not clear whether it is improved productive efficiency or increased saving that affects growth. The explanation based on fixed market costs proposed in the introduction presumes that:

1. growth and financial development have a long-run bi-directional positive relation,
2. financial market extensions can have persistent short-run negative effects on growth, and accumulated growth can cause financial market extensions.
3. the positive effect of financial development on growth is mediated through productive efficiency by specialization rather than through savings.
4. indirect effects from financial development through saving might go in either direction but are more likely to be negative if people are risk averse.

A set of papers use time series to explicitly study the direction of causality. An early example is Jung [1986] who tests Granger causality in a sample of 56 countries. His findings are mixed and indicate a changing temporal relation over time. Demetriades and Hussein [1996], apply modern time series techniques to conduct causality tests on time series from 16 countries that are not highly industrialized and conclude that the relationship in most of the countries is bi-directional. Neusser and Kugler [1998][4] estimate co-integration relations between the financial sector share and total factor productivity (TFP), and study Granger causality. Their evidence favors long-run TFP growth being caused by the financial sector, with a feedback effect, but the causal structure varies across countries. One interpretation of this evidence is that the combination of the first and second conjecture is observed at different stages in different countries. The third conjecture is also supported by this last study.

Berthélemy and Varoudakis [1995, 1996] find convergence clubs in a 1960-1985 cross-section sample of 91 countries. Financial development is found to have a positive impact on growth contingent on a sufficient educational level. A high correlation between educational level and GDP level indicates a productivity feedback.De Gregorio and Guidotti [1995] report further empirical evidence pointing in that direction. They find that most of the growth effects of financial services are through the marginal productivity of capital and not through investment volume. They also find a positive relation to growth in a cross-section of about 100 countries from 1960 to 1985, but a negative relation in a sample of six-year panel data for 12 Latin American countries between 1950-1985. Although they explain the negative short-run relation by financial repression, it could also be interpreted as a short-run negative growth effect from financial market extensions, thus supporting the second conjecture.

Atje and Jovanovic [1993] find effects on growth from stock markets in a broad sample cross-country regression based on Greenwood and Jovanovic [1990], but fail to find this for bank lending. A partial explanation might be found in Jappelli and Pagano [1994] who find that household credit has a negative influence in mature countries, and attribute this to the liquidity constraints in less developed countries, which force higher saving for the young generation. This would also help explain the Fernandez and Galetovic [1994] result and be consistent with the fourth conjecture in developed countries.

Some direct firm evidence points to scale effects in the financial sector that would be consistent with fixed market costs. Sussman and Zeira [1995] and Harrison et al.[1999] find in U.S. cross-state data that banking costs are lower where income per capita is higher and financial centers are geographically closer. In Saint-Paul [1992b] a circular localization model based on the same idea of fixed financial costs as Saint-Paul [1992a] explains dual economies by the segmentation of financial markets, making it hard to diversify financially in some branches or regions of the economy. Empirical evidence from Japan (over branches) and Italy (over regions) is consistent with this.

Finally, regime shifts in the relations between saving, growth and financial sector shares are reported in Lindh and Lindström [1997] for the G-7 economies. The structural breaks can be tied to major institutional changes in the financial markets and regulations. Jacobson et al. [1998] apply Markov-switching regressions to U.S. data 1948-1996 and find evidence of several regime shifts that also can be connected to major institutional changes in the financial system. Short-run Granger causality analysis indicates that the financial share Granger causes saving with a negative effect but is otherwise inconclusive. All three variables combine to predict regime shifts, indicating that this is endogenously determined.

A MODEL OF FINANCIAL MARKET INTERACTION WITH GROWTH AND SAVING

The model modifies the conventional two-period overlapping generations (OLG) model in several dimensions. Agents are spatially segregated. Investing in other locations requires an additional fixed cost. Technological choice is endogenous and made by agents, acting as entrepreneurs. Following the tradition, labor supply is assumed to be exogenous and normalized to unity at each location, with the additional restriction that it is only supplied to the local market. This allows us to view each local set of agents as a representative individual.

Insurance against the uncertainty of old age income is achieved either by flexible, less specialized technologies at the cost of lower productivity, or by diversifying investment to other locations at a fixed cost. Extensions of the financial markets are discrete events. Specialization, on the other hand, is not immediate, due to convex adjustment costs, since time-consuming investment, research, and learning are necessary.

Formal Specification

Local production at location i per labor unit is x_i and k_i is the average capital services per labor unit. Without loss of generality it is assumed that all capital is consumed over the production period. The local technology is affected by positive externalities from k_i such that average production becomes

$$x_i = A_i k_i^{1-b} \text{ and } A_i = B_i k_i^b \tag{1}$$

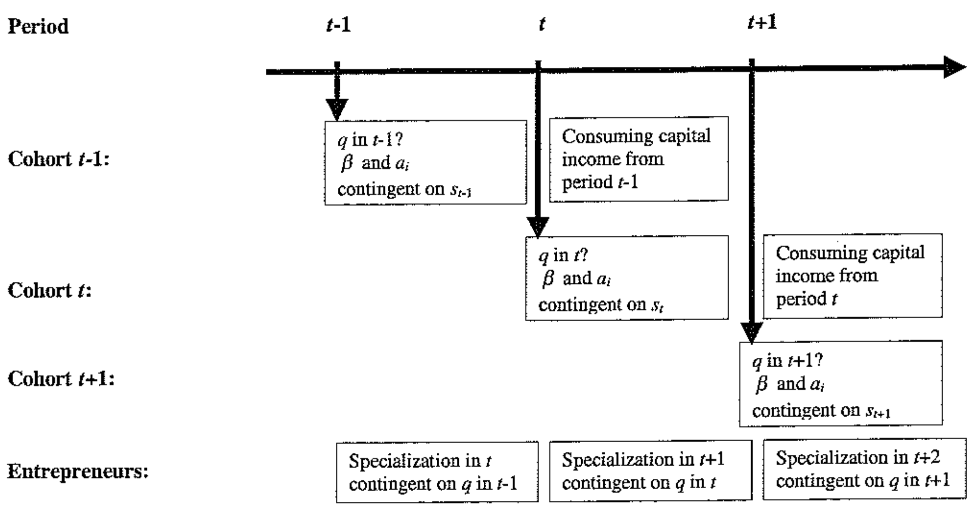
Since aggregate local production $B_i k_i$ is linear in k_i there are non-decreasing returns to capital.[5] The technology factor B_i is a random variable with a distribution conditional on the degree of specialization—a continuous index s_i. The choice of technology is restricted, such that the expected value of B_i and its variance both increase strictly with s_i.[6] The B_i in different locations are independently distributed and identical for given s_i. Specialization will not be more explicitly modeled so this index could also be more broadly interpreted as any increase in expected productivity that is accompanied by increasing risk.

Financial markets link different locations by providing mutual diversification opportunities. The analysis here is conducted from the point of view of one specific location taking the rest of the world as given. At each time there are a number of possible connections to other locations, providing access to already established financial markets. The alternative extensions are ordered by a discrete index q. This index measures the extent of the financial markets, such that, *ceteris paribus*, the expected utility of a given investment is increasing in the index q.[7]

It takes at least one period to start changing the degree of specialization. Agents therefore take the current distributions, as well as the realization of their stochastic labor income, as given. A competitive labor market ensures that labor income y equals the marginal product of labor $bA_i k^{1-b}$ (skipping indices henceforth when unnecessary). Consequently capital income is $(1-b)x = (1-b)Bk$, which is consumed in the retirement period.

Let F_q be total financial costs at financial extension q, such that $F_{q+1} > F_q$ by some fixed amount. Let δ be the discount factor and β the saving share of disposable income after financial costs, $y^d = y - F_q$. Young agents maximize expected utility over their lifetimes by choosing the share of disposable income to save and by adjusting their portfolios of investment shares a_i in different locations.

FIGURE 2
Schematic Diagram of the Sequence of Decisions in the Model



(2) $$\max_{\beta,a_i} u\left((1-\beta)y_t^d\right) + \delta E u\left((1-b)\sum_i B_i a_i \beta y_t^d\right) \text{ subject to } \sum_i a_i = 1,$$

where c is consumption, and $u(c)$ is a constant relative risk-aversion (CRRA) function.

(3) $$u(c)=\frac{c^{1-p}-1}{1-p}, \quad p>0 \text{ and } p\neq 1.$$

If $p = 1$ then $u(c) = \log c$. The distribution of B_i is Gaussian, or has finite support, so a unique maximum of expected utility of capital returns, $Eu((1-b)A^{-b}$, with respect to s, exists. A competitive capital market equalizes returns within locations. Entrepreneurs choose both capital intensity and specialization, but the latter is only gradually adjusted to its optimal value. This choice is not formally modeled since free entry for entrepreneurs will ensure that s converges to its expected utility-maximizing value. The competitive factor markets ensure that the local realization of B_i is the same for all local entrepreneurs, which is why we do not have to distinguish between a firm's capital intensity and the average capital intensity.

If there were no financial costs, risk-averse agents would prefer to insure against risk by spreading investments over as many locations as possible, (i.e., integrating the financial markets maximally). The presence of a fixed cost implies that financial diversification will only be taken to the point where the incremental expected utility offsets the decline in utility from the incremental cost. The young agent must therefore also decide at which level of financial diversification his expected utility will be maximized.

This choice is made by comparing maximal utility in equation (2) for two financial extensions, while taking current specialization as given. The sequence and timing of decisions in the model are summarized in Figure 2.

Properties of the Equilibrium

Conditional on q and s, the first order necessary maximization conditions for equation (2) require that marginal utilities between periods are equalized by choosing β

(4) $$y_t^d u'(c_t) = \delta \frac{\partial Eu(c_{t+1})}{\partial \beta}$$

and, simultaneously, that marginal utilities of each local investment are equalized by the portfolio choice

(5) $$\delta \frac{\partial Eu(c_{t+1})}{\partial a_i} = m \quad \text{for} \quad i = 1,\dots,n_q,$$

where m is the shadow value of each investment unit. The first-order condition in equation (4) with respect to β can be rewritten, using the CRRA specification (3) and dividing by $(y_t^d)^{1-p}$, as

(6) $$(1-\beta)^{-p} = \delta(1-b)^{1-p}\beta^{-p}\int_{\Omega_q}\left(\sum_i B_i a_i\right)^{1-p} d\Phi_q,$$

where Φ_q is the joint distribution function of the random production returns B_i at the locations encompassed by the current financial markets, and Ω_q is its support.[8] The LHS increases monotonically with β, while the RHS decreases monotonically in β.

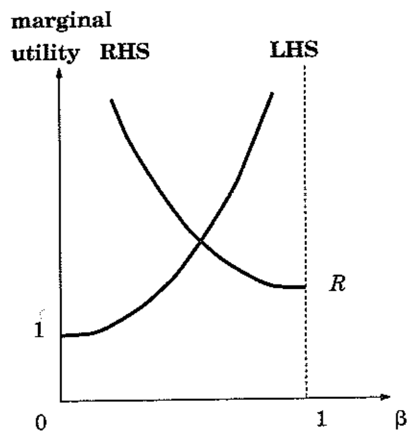
For $\beta = 0$ the LHS of equation (6) is obviously positive but finite. The RHS goes to infinity. For $\beta = 1$ the opposite holds, i.e., the LHS goes to infinity, while the RHS is finite and positive. It is clear that the two schedules must intersect as in Figure 3. Existence and uniqueness of an optimum is thereby proved.

Define for short $R = \delta(1-b)^{1-p}\int_{\Omega_q}\left(\sum_i B_i a_i\right)^{1-p} d\Phi_q$ which is independent of β. R is monotonically related to the expected utility of a unit investment but not exactly the same. We can then write (6) more succinctly as

(7) $$(1-\beta)^{1-p} = \beta^{-p}R.$$

In Figure 3 the LHS schedule is therefore fixed and the RHS schedule shifts up (down) as R increases (decreases).

FIGURE 3
Schedules of the Left and Right-Hand Sides of the First Order Conditions for Optimal Saving and Investment Decisions



Financial Extensions and Specialization

With the degree of specialization given, there are three effects of an extension of the financial regime. First, it decreases variance in expected old age income; second, it decreases disposable labor income; third, the optimal β changes as the asset portfolio is reallocated. To analyze how β changes we only need to know how R changes.

Expand the integral in R in a second order Taylor approximation around the expected value, $\mu = E(\sum_i B_i a_i)$.

$$E\left(\left(\sum_i B_i a_i\right)^{1-p}\right) \approx \mu^{1-p} - \frac{1}{2}(1-p)p\mu^{-p-1}Var\left(\sum_i B_i a_i\right) \tag{8}$$

If expected returns are constant over the two regimes it is obvious that R decreases (increases) as variance is decreased if $p > 1$ ($p < 1$) and the optimal savings share will decrease (increase) correspondingly. With low risk aversion the agent will shift disposable labor income to investment as risk decreases to increase future consumption. With higher risk aversion he will prefer to shift disposable income to current consumption instead. A slightly more complicated argument in the Appendix shows that this conclusion holds also when expected returns change.

To the agent an extension becomes favorable as the optimal expected utility in the current financial regime becomes less than in the alternative. Eliminating constants on both sides, the equilibrium at extension q becomes sub-optimal as

$$\frac{c_{t,q}^{1-p}}{1-p} + \delta \int_{\Omega_q} \frac{c_{t+1,q}^{1-p}}{1-p} d\Phi_q < \frac{c_{t,q+1}^{1-p}}{1-p} + \delta \int_{\Omega_{q+1}} \frac{c_{t+1,q+1}^{1-p}}{1-p} d\Phi_{q+1} . \tag{9}$$

The first order conditions in equation (4) imply that

$$\frac{c_t^{1-p}}{1-\beta} = \delta \int_{\Omega} \frac{c_{t+1}^{1-p}}{\beta} d\Phi \tag{10}$$

Using this and multiplying both sides with $1 - p$ we get, when $p > 1$

$$\frac{c_{t,q}^{1-p}}{1-\beta_q} < \frac{c_{t,q+1}^{1-p}}{1-\beta_{q+1}} . \tag{11}$$

If $p < 1$ the inequality is reversed. Since $c = (1 - \beta)(y - F)$ this implies

$$\left(\frac{y-F_{q+1}}{y-F_q}\right)^{1-p} < \left(\frac{1-\beta_{q+1}}{1-\beta_q}\right)^{p} , \tag{12}$$

This condition will eventually hold, if disposable income growth is strictly positive. Since $p > 1$ the savings share $\beta_q > \beta_{q+1}$, so the RHS in equation (12) is greater than unity and the LHS converges to unity from above as y grows. Reversing inequalities when $p < 1$ the condition will again hold eventually. Then $\beta_q < \beta_{q+1}$, so the RHS in equation (12) is less than one while the LHS goes to one from below.

Total savings, βy^d, will obviously fall when $p > 1$ as the financial market is extended. For $p < 1$ total savings will decrease in some cases, too, since disposable labor income decreases.[9]

As entrepreneurs adapt s to the new financial regime, risk increases in production, but is balanced by increases in the expected value of B. Entrepreneurs will only change s to increase the expected utility of agents, but, as above, the behavior of R depends on risk-aversion. It is shown in the Appendix that the R again decreases (increases) when $p > 1$ ($p < 1$). So when specialization has adapted, the savings share will be even lower (higher).

If F_q was increasing with output or investment, specialization might ultimately de-stabilize the new financial extension and lead to a contraction of the financial markets. Depending on the rate of growth, this may even give rise to cyclical behavior. Such cost behavior could for example arise if costs are tied to the volume of trading, as they often are.

However, in the real world we would also expect some offsetting factors, (e.g., economies of scale in information handling, a downward drift in financial costs due to improved transaction technologies, and external effects from neighboring financial markets).

TABLE 2
Summary of the Expected Effects in the Transition to an Extended Financial Regime and the Following Increase in Specialization if *F/y* Decreases

Transition	p>1	p=1	p<1	Specialization	p>1	p	p<1
β	–	0	+	β	–	0	+
$\beta y^d/y$	–	–	?	$\beta y^d/y$	?	?	?
g	–	–	?	g	+	+	+
F_q/y	+	+	+	F_q/y	–	–	–

Growth Rates

For a tractable analysis of growth rates we simplify further in this subsection by assuming that the B_i are identically distributed over all locations encompassed by the current financial markets. This implies that investment in local production equals the total investment of local agents.[10] Then the growth rate of output is

(13) $$g_x = \frac{x_{t+1}}{x_t} - 1 = \frac{Bk_{t+1}}{x_t} - 1 = \frac{B\beta y_t^d}{x_t} - 1 = B\beta\left(b - \frac{F_q}{x_t}\right) - 1 .$$

The growth rate of local disposable income (for the young) is higher

(14) $$g_y = \frac{y_{t+1}^d}{y_t^d} - 1 = \frac{bx_{t+1} - F_q}{y_t^d} - 1 = B\beta b - \frac{F_q}{y_t^d} - 1$$

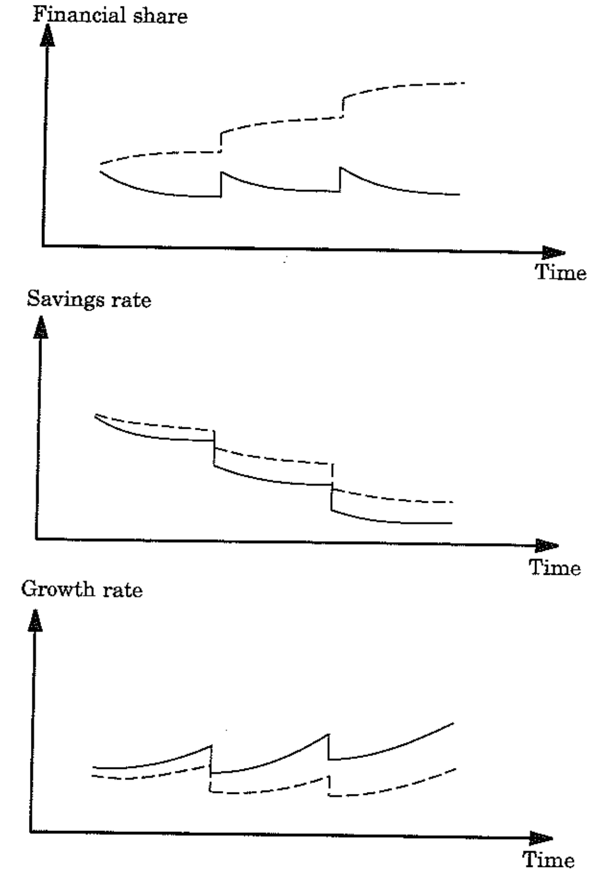
as can be seen by taking the difference

(15) $$g_x - g_y = \frac{F_q}{y_t^d} - \frac{B\beta F_q}{x_t} = F_q \frac{x_t - B\beta y_t^d}{x_t y_t^d} = \frac{-F_q}{y_t^d} g_x .$$

With positive growth in y, g_y converges to g_x which in turn converges to $bB\beta - 1$, if financial costs decrease relative to income. The financial regime associated with a higher q will on average have an initial slump in the growth rate since βy_t^d decreases for empirically plausible values of the risk aversion parameter. As entrepreneurs adjust specialization to the new regime, the expected value of B will increase, and hence the growth rate also decreases while β decreases further. Recall that a decreasing saving rate refers to y^d. Since the relative leakage through financial costs decreases, the effect on savings as a share of y is *a priori* ambiguous. Expected utility increases with specialization and variance increases. The increase in risk is balanced by an increase in the expected growth rate. Of course, the higher the risk aversion, the more sluggish will the increase in the growth rate be.

Whether growth rates increase relative to the previous equilibrium depends on the parameters. The pattern of growth to expect from the model would be a saw-tooth diagram of the growth rate, collapsing as a new financial regime was introduced and then slowly recovering again. The financial cost share would be expected to show

FIGURE 4
Two Possible Patterns Implied by the Model for Respectively the Financial Share, the Saving Share of Disposable Income, and the Growth Rate



sharp increases at transitions. Whether this share tends to grow or decrease between transitions would depend on exactly how costs are related to the scale of production. In any case the relationship between growth and the financial share would not be linear over time. Savings as a share of gross product would decrease at the transition and is indeterminate between transitions.

Table 2 summarizes the effects predicted by the model for different values of the risk-aversion parameter. These predictions are, however, contingent on the assumption that F/y decreases in the period following an extension. In Figure 4 the dashed lines presume that financial costs increase with production fast enough to dominate the scale economies generated by the fixed costs. This gives an impression of the differences. These graphs are not formally derived from the model but illustrate possible patterns that are consistent with the model.

In the real world there are interactions between countries, changes in institutional and political regimes, and so forth, that will cause patterns to differ over time. Moreover, there are adjustment costs, which are likely to cause considerable delays in reactions, and so on. Referring back to Figure 1 it would be natural to interpret the expansionary periods as corresponding to the jumps in Figure 4, where the financial extension is drawn out over a longer period. The message of Figure 4 should thus not be interpreted too literally; the empirical prediction of the model is not that we will observe regular patterns like those in the graphs. What we would expect to see if the model has empirical relevance is coordinated regime switches in the behavior of saving, growth and the financial cost share. That has indeed been observed [Jacobson et al., 1998; Lindh and Lindström, 1997].

The question is: are fixed financial costs really that important? Harrison et al. [1999] estimate that banking costs in the United States rose from 4 percent of extended credit in 1965 to 6-7 percent in the 1990s. Taken literally in the model this implies an increased leakage in the transformation of savings to capital and should slow capital accumulation down to the same extent. In the simple growth model here, this is immediately translated into growth rates. Thus, a slowdown on the order of 2 percent of the growth rate since the 1960s is quite well in line with the evidence on financial costs.

CONCLUDING COMMENTS

The OLG model of discontinuous financial development analyzed in this paper implies a non-linear relation between growth and the financial sector share. Each new financial market formation is associated with fixed costs leading to a temporary growth slump. In the long run the improved risk sharing provided by the new markets induce more risky but also more productive specialization that raises growth rates again. However, the fixed costs constitute a leakage in the transformation of saving into investment. If agents have iso-elastic risk preferences with a risk-aversion parameter above unity, precautionary saving will fall, too. Thus, even in the long run saving rates will be depressed. When the scale of production has grown enough to offset the cost of the next financial extension, the cycle starts over again.

The stylized model reconciles the observed short-run slow-down in growth rates when financial markets expand with the stylized fact that financial development enhances long-run growth, as well as providing an explanation for the decreasing saving rates that can be empirically linked to financial development. The decrease in saving rates implies that the positive growth effects of further financial extensions will exhibit diminishing returns, because the increases in productive efficiency are offset both by decreases in saving rates and by an increased cost leakage that slows down capital accumulation. Thus, financial development should prove most beneficial to the least developed countries.

The interpretation of financial markets in the model is entirely in terms of diversification of risk over investment opportunities, (i.e. as stock markets). But financial pooling has other benefits, like transactional and informational efficiency. It remains to explore how such mechanisms would interact with those modeled here, but it seems unlikely that this would radically change the conclusions. The basic idea of the model could, in principle, be extended to any form of financial insurance, including social welfare and other institutions for income insurance.

A natural extension would be to consider insurance against variability in labor income by a social welfare state or other institutions like long-term labor contracts. One would expect that differences in cost structure between different insurance devices,(i.e. social institutions), would become crucial to long-term growth.

Finally, it could be noted that this kind of model is a natural complement to the literature on technological shifts. As David [1991] shows very convincingly with regard to the electric dynamo there is a quite considerable lag of 20 to 30 years before new general purpose technologies actually starts to yield a productivity payoff. This is because new intermediate components, new organizations, and new distribution channels need to be developed and not least the labor force needs to be trained, before the potential in the new technology can be realized. Many believe that the so called "new economy" is a sign of the computer technology finally starting to yield productivity gains, which previously have been hard to find. A quite extensive literature—both theoretical and empirical—now studies such technology shifts, (see Helpman [1998] for a sample). This relates to this paper in that the initial phase of development and training calls for quite extensive investment to replace old technology, thus producing a demand for increased financial intermediation. In Eriksson and Lindh [2000] a formal model of this process is shown to produce patterns very similar to those implied by the financial development model in this paper.

APPENDIX

In the section entitled, "Financial extensions and specialization" it was asserted, that the saving rate increases (decreases) as the risk-aversion parameter is below (above) unity and also when expected returns change. This assertion is formally proved here.

Proposition:

a) The saving rate β increases (decreases) at the transition to a more extended financial market when the risk-aversion parameter $p < 1$ ($p > 1$). If $p = 1$the saving rate is constant.

b) The saving rate β increases (decreases) as production becomes more specialized when the risk-aversion parameter $p < 1$ ($p > 1$). If $p = 1$ the savings rate is constant.

Proof:

a) First note that equation (7) can be written $(\beta/(1-\beta))^p = R$ Taking the differential confirms that β increases monotonically with R. Second, note that the difference in optimal expected utility at a transition must be positive:

$$\text{(A.1)} \quad \frac{\int\limits_{\Omega_{q+1}} ((1-b)\beta_{q+1} y_{q+1}^{d} \sum_{i=1}^{q+1} a_i B_i)^{1-p} \, d\Phi_{q+1} - \int\limits_{\Omega_q} ((1-b)\beta_q y_q^{d} \sum_{i=1}^{q} a_i B_i)^{1-p} \, d\Phi_q}{1-p} > 0$$

$0 < p < 1$: Equation (A.1) implies that $(\beta_{q+1} y_{q+1}^{d})^{1-p} R_{q+1} > (\beta_q y_q^{d})^{1-p} R_q$. Hence if $R_{q+1} < R_q$ it follows that $\beta_{q+1} y_{q+1}^{d} > \beta_q y_q^{d}$. Since $y_q^d > y_{q+1}^d \Rightarrow \beta_{q+1} > \beta_q$ thus contradicting that β increases monotonically with R. Therefore it must be that $R_{q+1} < R_q$ and $\beta_{q+1} > \beta_q$.

(*ii*) $p > 1$: Equation (A.1) in this case implies that $(\beta_{q+1} y_{q+1}^{d})^{1-p} R_{q+1} < (\beta_q y_q^{d})^{1-p} R_q$. By equation (7) we get

$$\text{(A.2)} \quad \frac{\beta_{q+1}^{1-p} R_{q+1}}{\beta_q^{1-p} R_q} < \left(\frac{y_q^d}{y_{q+1}^d} \right)^{1-p} < 1 \Rightarrow \frac{\beta_{q+1} R_{q+1}}{\beta_q R_q} < \frac{\beta_{q+1}^p}{\beta_q^p} = \frac{(1-\beta_{q+1})^p R_{q+1}}{(1-\beta_q)^p R_q}$$

so we can divide away the Rs. Since $\beta_{q+1} < \beta_q \Rightarrow 1 - \beta_q > 1 - \beta_{q+1}$ we obtain a contradiction and can conclude that $\beta_{q+1} < \beta_q$.

$p = 1$: Then $u(c) = \log c$ and the first order condition w.r.t. β reduces to $\beta/(1-\beta) = \delta$ so the savings rate will not depend on financial extension.

b) In this case no fixed cost need be considered, so it is sufficient to review the case when specialization increases between two periods with given disposable income. If expected utility is increased it must hold that

$$\text{(A.3)} \quad \frac{\int\limits_{\Omega_{s_{t+1}}} (\beta_{s_{t+1}} \sum_{i=1}^{q} a_i B_i)^{1-p} \, d\Phi_{s_{t+1}} - \int\limits_{\Omega_{s_t}} (\beta_{s_t} \sum_{i=1}^{q} a_i B_i)^{1-p} \, d\Phi_{s_{t+1}}}{1-p} > 0$$

(*i*) $0 < p < 1$: Essentially the same argument as in *a*) (*i*) remains true only with the modification that we do not need the argument that disposable income has decreased.

(*ii*) $p > 1$: Again we proceed as in case *a*) but in this case the middle part of the first inequality in equation (A.2) is not needed. Alternatively note that the change in the distribution as specialization increases will, by assumption, cause more density to be concentrated at higher outcomes on average. The fact that the integrand is decreasing in outcome will then directly imply that $R_{s_{t+1}} < R_{s_t}$ and $\beta_{s_{t+1}} < \beta_{s_t}$.

(*iii*) $p = 1$: Exactly the same argument as in *a*). **QED**

NOTES

I thank Clas Eriksson, Tomas Lindström and Henry Ohlsson for useful suggestions and comments. Comments from seminar participants at Uppsala University and a number of conferences have also been useful in improving the paper. I gratefully acknowledge research funding by the American Council of Learned Societies and the Swedish Council for Research in the Humanities and Social Sciences.

1. The data are calculated from the National Income and Product Accounts of the United States (1929-1947 from NIPA 1993 and 1948-1996 from NIPA 1997). The financial share is obtained by dividing row 18 (financial corporate GDP) by row 1 (total corporate GDP) in current value estimates in table 1.16. The growth rate is log (x_t/x_{t-1}) where x is row 36 (real non-financial corporate GDP) deflated by a chain index back to 1948. For the years 1929-1947 real value estimates of this variable were not available. The current value estimates (row 19) were therefore deflated by the implicit price deflator for the business sector (row 2 in Table 7.14). These two real series were then ratio linked in 1948. The two deflators for overlapping years differ in the late 1970s and early 1980s, but otherwise are very similar. The saving share is gross saving divided by GDP, Table 5.1, row 1 divided by Table 1.1, row 1.
2. The Great Depression, World War II and the Korean War may seriously distort data in the first two periods. Before 1958 data are not necessarily consistent with later definitions and deflators especially are less reliable. The evidence for the two later periods should therefore be given more consideration than the two previous periods.
3. Most cross-section growth correlations have turned out to be very sensitive to changes in the set of control variables—as emphasized in a seminal paper by Levine and Renelt [1992]. It is therefore important to point out that Levine and Zervos [1993, 1996] find the correlation between growth and several financial development indicators (*inter alia* risk price measures) robust in this sense.
4. They use financial sector GDP to measure financial development, as does Kuznets [1971, 107] He notes an unusually clear relationship between average income and the proportion of the finance, insurance, and real estate sector in a cross-section of 27 countries in 1958.
5. Besides ensuring a simple endogenous growth mechanism, this has the convenient property that all growth paths are balanced. It also implies that higher saving boosts the growth rate permanently. See Rebelo [1991] for more about the linear capital model as a representative of the class of endogenous growth models with convex technology.
6. These two statistics are taken to be sufficient for agents to rank the expected utility of local technologies. Commonly used ranking criteria such as mean-preserving spread and second order stochastic dominance implies increasing variance with risk.
7. The set of possible connections can obviously be partially ordered in terms of costs and diversification opportunities. The trade off of costs against diversification will under the assumptions below be independent of the scale of investment, implying that alternatives can in fact be linearly ordered.
8. By rewriting the expectation integral explicitly: $\int u\left((1-b)\sum_i B_i a_i \beta y_t^d\right) d\Phi_q$ differentiating w.r.t. β and a_i, and using the linearity properties of the integral operator, it follows that $\frac{\partial Eu}{\partial \beta} = \sum_i \frac{a_i}{\beta} \frac{\partial Eu}{\partial a_i} = \frac{m}{\beta\delta}$. From equation (6) it is clear that m is independent of β only in the special case when $p=1$.
9. It can be shown that it holds in the neighborhood of $p=1$, which is obvious, but also as $p \rightarrow 0$. and agents become risk-neutral. To exactly delimit the potential set in between, where it might not hold, the distributions of B_i and the cost F need to be specified. Since p empirically is found to be quite high, this is of limited interest. Kandel and Stambaugh [1991] give an overview of estimates. They argue on theoretical grounds that the relative risk aversion may be as high as 20 or above. There is no consensus on any exact number, but most estimates, both from micro and macro data, are well above one. In the context of a CRRA utility function the reciprocal value of the relative risk-aversion parameter is the intertemporal elasticity of substitution, which thus should be well below unity. Hall [1988] for example finds values of this elasticity below 0.1.

10. The more general setting above would not qualitatively affect the argument below, but the growth rate would depend on the weighted averages of parameters and variables over the locations encompassed by the current financial market.

REFERENCES

Acemoglu, D. and Zilibotti, F. Was Prometheus Unbound by Chance? Risk, Diversification, and Growth. *Journal of Political Economy*, August 1997, 709-51.

Atje, R. and Jovanovic, B. Stock Markets and Development. *European Economic Review*, April 1993, 632-40.

Bencivenga, V. R. and Smith, B. Financial Intermediation and Endogenous Growth. *Review of Economic Studies*, April 1991, 195-209.

Berthélemy, J.C. and Varoudakis, A. Thresholds in Financial Development and Economic Growth. *Manchester School of Economic and Social Studies*, Supplement 1995, 70-84.

________. **and** ________. Economic Growth, Convergence Clubs and the Role of Financial Development. *Oxford Economic Papers*, April 1996, 300-28.

David, P.A. Computer and the Dynamo: The Modern Productivity Paradox in a Not-Too-Distant Mirror, in *Technology and Productivity: The Challenge for Economic Policy*. Paris: OECD, 1991, 315-47.

De Gregorio, J. and Guidotti, P. E. Financial Development and Economic Growth. *World Development*, March 1995, 433-48.

Demetriades, P. O., and Hussein, K. A. Does Financial Development Cause Economic Growth? Time Series Evidence from 16 Countries. *Journal of Development Economics*, December 1996, 385-409.

Devereux, M. B. and Smith, G. W. International Risk Sharing and Economic Growth. *International Economic Review*, August 1994, 535-50.

Eriksson, C. and Lindh, T. Growth Cycles with Technology Shifts and Externalities. *Economic Modelling*, January 2000, 139-70.

Fernandez, D. G. and Galetovic, A. Schumpeter Might Be Right—But Why? Explaining the Relation Between Finance, Development, and Growth. Manuscript, Johns Hopkins University, Baltimore, 1994.

Goldsmith, R.W. *Financial Intermediaries in the American Economy Since 1900*. NBER, Studies in Capital Formation and Financing No. 3. Princeton: Princeton University Press, 1958.

________. *Financial Structure and Development*. New Haven: Yale University Press, 1969.

Greenwood, J. and Jovanovic, B. Financial Development, Growth, and the Distribution of Income. *Journal of Political Economy*, October 1990, 1076-1107.

Hall, R.E. Intertemporal Substitution in Consumption. *Journal of Political Economy*, April 1988, 339-57.

Harrison, P.; Sussman, O. and Zeira, J. Finance and Growth: Theory and New Evidence. FEDS Working Paper No. 99-35, July 1999, Federal Reserve Bank of Philadelphia.

Helpman, E. (ed.) *General Purpose Technologies and Economic Growth*. Cambridge, Massachusetts: MIT Press, 1998.

Jappelli, T. and Pagano, M. Saving, Growth, and Liquidity Constraints. *Quarterly Journal of Economics*, February 1994, 83-109.

Jacobson, T., Lindh, T., and Warne, A. Growth, Savings, Financial Markets and Markov Switching Regimes. Sveriges Riksbank Working Paper Series, No. 69, September 1998, Central Bank of Sweden, Stockholm.

Jung, W.S. Financial Development and Economic Growth: International Evidence, *Economic Development and Cultural Change*, January 1986, 333-46.

Kandel, S. and Stambaugh, R.F. Asset Returns and Intertemporal Preferences. *Journal of Monetary Economics*, February 1991, 39-71.

Khan, A. Financial Development and Economic Growth. FEDS Working Paper No. 99-11, September 1999, Federal Reserve Bank of Philadelphia.

King, R. and Levine, R. Finance, Entrepreneurship, and Growth: Theory and Evidence. *Journal of Monetary Economics*, December 1993a, 513-42.

________. **and** ________. Finance and Growth: Schumpeter Might Be Right. *Quarterly Journal of Economics*, August 1993b, 717-37.

Kuznets, S.S. *Economic Growth of Nations, Total Output and Production Structure*. Cambridge, Massachusetts: The Belknap Press of Harvard University Press, 1971.

Levine, R. Financial Development and Economic Growth: Views and Agenda. *Journal of Economic Literature*, June 1997, 688-726.

Levine, R. and Renelt, D. A Sensitivity Analysis of Cross-Country Growth Regressions, *American Economic Review*, September 1992, 942-63.

Levine, R. and Zervos, S.J. What We Have Learned about Policy and Growth from Cross-Country Regressions? *American Economic Review*, May 1993, Papers and Proceedings, 426-30.

________. **and** ________. Stock Market Development and Long-Run Growth. *World Bank Economic Review*, May 1996, 323-39.

Lindh, T. and Lindström, T. Structural Breaks in the Financial Sector Share, Savings and GDP Growth in OECD: 1960-1992. *Applied Economics*, January 1997, 33-41.

Neusser, K. and Kugler, M. Manufacturing Growth and Financial Development: Evidence from OECD Countries. *Review of Economics and Statistics*, November 1998, 638-46.

Obstfeld, M. Risk-Taking, Global Diversification, and Growth. *American Economic Review*, December 1994, 1310-29.

Pagano, M. Financial Markets and Growth: An Overview. *European Economic Review*, April 1993, 613-22.

Rebelo, S. Long-Run Policy Analysis and Long-Run Growth. *Journal of Political Economy*, June 1991, 500-21.

Saint-Paul, G. Technological Choice, Financial Markets and Economic Development. *European Economic Review*, May 1992a, 763-81.

________. Technological Dualism, Incomplete Financial Markets and Economic Development. *Journal of International Trade & Economic Development*, 1992b, 12-26.

Sussman, O., and Zeira, J. Banking and Development, Discussion Paper No. 1127, Centre for Economic Policy Research, London, 1995.

U.S. Department of Commerce. *National Income and Product Accounts of the United States, I and II*, data released January 1993. Washington D.C.: Bureau of Economic Analysis.

________. *National Income and Product Accounts of the United States, I and II*, data released April 1997. Washington D.C.: Bureau of Economic Analysis.