No. 1127

BANKING AND DEVELOPMENT

Oren Sussman and Joseph Zeira

FINANCIAL ECONOMICS AND INTERNATIONAL MACROECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

BANKING AND DEVELOPMENT

Oren Sussman and Joseph Zeira

Discussion Paper No. 1127
February 1995

Centre for Economic Policy Research
25–28 Old Burlington Street
London W1X 1LB
Tel: (44 71) 734 9110

This Discussion Paper is issued under the auspices of the Centre's research programmes in **Financial Economics** and **International Macroeconomics**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

CEPR, 25–28 Old Burlington Street, London W1X 1LB. Tel: 44 71 734 9110. Fax: 44 71 734 8760.

CEPR Discussion Paper No. 1127

February 1995

ABSTRACT

Banking and Development*

This paper reformulates the well known financial development conjecture (FDC) and supplies some new empirical evidence in its favour. The financial development conjecture, namely, that there exist strong feedback effects between real and financial development, is described in this paper by use of the cost of financial intermediation. The theoretical part of the paper describes how specialization of banks can lead to such feedback effects, which work through the cost of financial intermediation. In the empirical part of the paper we use US cross-state data from banks' income statements to show that the cost of banking is negatively related with the level of real economic development.

JEL Classification: E44, G21, O16
Keywords: financial development, banks, economic growth

Oren Sussman and Joseph Zeira
Department of Economics
Hebrew University of Jerusalem
Mount Scopus
Jerusalem 91905
ISRAEL
Tel: (972 2) 883 075/256

*This paper is produced as part of a CEPR research programme on *Market Integration, Regionalism and the Global Economy,* supported by a grant from the Ford Foundation (no. 920–1265). We would like to thank Alex Cukierman, Zvi Eckstein, Stanley Fischer, Saul Lach and Nathan Sussman for helpful comments on an earlier draft, Yoel Naveh for very good research assistance, and the Chilewich foundation for financial support.

Submitted 17 January 1995

NON-TECHNICAL SUMMARY

This paper reformulates the well known financial development conjecture and presents new evidence in its favour. The financial development conjecture claims that there exist strong feedback effects between real economic development and financial development.

The theoretical part of the model presents an information based model of a monopolistically competitive banking system. In this model financial intermediaries (banks) are formed in order to economize on the costs of monitoring *ex-post* returns of defaulting entrepreneurs. Banks are therefore modelled as delegated monitors of borrowers, following the literature on financial intermediation by Townsend (1979), Diamond (1984) and others. We then define the cost of financial intermediation as the costs of transferring one unit of loan from an ultimate lender to an ultimate borrower (a firm, or a project) as a result of the existence of monitoring costs. Banks and firms are then modelled in a spatial environment, where the distance between a bank and a borrowing firm affects the cost of monitoring. We assume that costs increase with distance. This allows for specialization by banks, which generates a feedback effect between the real and the financial sector in the following way. As the economy grows, more capital is invested and channelled through banks. This raises banks' profits and thus induces the entry of more banks. This entry increases specialization as each bank deals with a closer set of borrowers. Hence the average distance between banks and firms is reduced and that reduces the cost of financial intermediation. As a result, funds become less expensive, since the cost of financial intermediation is part of the wedge between lenders' and borrowers' interest rates, and hence investment and growth increase. This feedback effect therefore enables the financial sector to magnify effects of various shocks on the economy. We call this feedback effect the specialization effect.

This specialization effect is accompanied by another effect that works in the opposite direction, which we call the wage effect. Growth in the real sector raises wages and that raises the cost of monitoring, and as a result, the cost of financial intermediation. The model presents the two effects and also shows that their relative effectiveness depends on the way monitoring costs depend on distance, namely on the returns to specialization. Nevertheless, the issue of the relative strength of the two effects is an empirical one.

In the second part of the paper we use US cross-state aggregate banking data to provide some evidence on the relative strength of the specialization effect

and the wage cost effect. We show that the cost of financial intermediation is indeed lower in states with higher output per capita. This is a surprising result, because in general the costs of most services tend to rise along the development path, due to the wage cost effect (see Kravis and Lipsey (1983)). We interpret this finding as evidence in support of a strong specialization effect, namely in favour of the financial development conjecture.

The banking data we use present striking geographical characteristics, which give support to our choice of a spatial model. Thus, the north-east states have lower costs of banking relative to the western and southern states. Furthermore, distance from the nearest financial centre appears to be a variable with a significant effect on the cost of banking. Our results show that the costs of banking are strikingly high: 5–6% of banks extended credit on average. We also find that *ex-post* monitoring is quite high: 40% of debt written off the balance sheet.

Recently, there has been growing research, both theoretical and empirical, on the financial development conjecture. Our work can be differentiated from these works in two aspects. First, we formulate the financial development conjecture in terms of the cost of financial intermediation which becomes the link between the real and the financial sector. This variable is observable (although with some difficulty) from banking data. Thus, our model allows for a narrowing of the gap between theory and data in this area. Second, this paper differs in its use of US cross-state data, as opposed to international data. We believe that such data have important advantages. The data are produced under common definitions and reporting standards and with small cross-state differences in regulatory regimes.

Banking and Development

1. Introduction

This paper reformulates the well known financial development conjecture (FDC) and presents some new evidence in its favor. The financial development conjecture is that there are strong feedback effects between real economic development and financial development. In the theoretical sections of the paper we construct a model of financial intermediation, in which the feedback between the two sectors stems from banks' specialization. A key variable in the analysis is the cost of financial intermediation per unit of loan. This formulation enables us to bring the theory closer to observed variables. Indeed, in the empirical sections of the paper we use US cross state aggregate data on banks, to show that the cost of banking is negatively correlated with the level of real economic development. We view this as evidence in favor of the FDC.

The theoretical part of the paper presents an information based model of a monopolistically competitive banking system. It builds on recent theoretical developments, mainly Townsend's (1979) and Gale and Hellwig's (1985) theory of standard debt contracts, Diamond's (1984) theory of the delegated monitoring role of banks and Salop's (1979) circle model of spatial competition. In our model financial intermediation is a way to economize the costs of monitoring ex-post returns of defaulting projects.[1] A crucial assumption we add, is that monitoring costs are increasing with the distance between the bank and the project. This allows for banks' specialization, which generates the feedback effect in the following way. As the economy grows, more capital is channeled through banks, which raises their profits and induces entry of more banks. This entry reduces the average distance between bank and projects, increases specialization, and thus

[1]Obviously, the cost of financial intermediation includes other types of costs as well, such as ex-ante monitoring costs, etc. We concentrate in our model on ex-post monitoring for reasons of tractability, and because adding other costs does not alter the main results of the model.

reduces the cost of financial intermediation. This in turn raises investment and real growth. We then show that the financial sector magnifies, through this feedback, the effect of productivity shocks on output. Our model, therefore, indicates that financial systems account for some of the international differences in output per capita.[2]

Note, that the specialization effect described above, which reduces the cost of financial intermediation, is accompanied by a second effect, the wage cost effect. Growth raises wages, and thus raises the cost of financial intermediation. The two effects, therefore, operate in opposite directions. While their relative effectiveness is an empirical issue, the theoretical model shows that it depends on the returns to specialization, namely on the way intermediation costs depend on distance.

In the empirical part of the paper we use US banking data to provide some evidence on the relative strength of the specialization effect and the wage cost effect. We show that the cost of banking is indeed lower in more developed states. Thus, for example, the north-east provinces have lower costs of banking relative to the western and southern provinces. In a more formal regression analysis, output per person typically comes out significantly negative. In a sense, this is a surprising result, because in general, the costs of most services tend to rise along the development path, due to the wage cost effect (see Kravis and Lipsey, 1983). We interpret our findings, therefore, as evidence for a strong specialization effect, namely in favor of the FDC.

Our empirical findings also give additional support to our choice of the cost of financial intermediation as a key variable, as it appears to be a significant variable. Indeed, the cost of banking is found to be strikingly high: 5-6% (on average) of banks' extended credit. The cost of ex-post monitoring is also

[2]This is related to the issue raised by Lucas (1990).

estimated and is higher than anticipated: 40¢ on 1$ of debt written off the balance sheet. Note, that these results are of importance even beyond the context of the present paper, in areas such as business cycles, monetary economics or corporate finance.

The financial development conjecture was already raised and discussed during the 1950s and the 1960s. The major works from that period include Gurley and Shaw (1955), Cameron (1967), Goldsmith (1969), McKinnon (1973), Shaw (1973), and Kindleberger (1974). Our paper belongs to a recent new wave of research on this issue, which differs from former work in being more formal and in stressing the financial intermediating role of banks, rather than their payment functions. This line of research includes Jovanovic and Greenwood (1989), Sussman (1993)[3], Atje and Jovanovic (1992), De la Fuente and Marin (1993), Zilliboti (1993), and is surveyed by Pagano (1991).[4]

Within the new research on the FDC, this paper concentrates on a feedback channel, which is as directly observable as possible, namely the cost of financial intermediation by banks. Thus, it tries to narrow down the substantial gap between theory and evidence in this area. This paper also differs from others in its use of US data, instead of international comparisons.[5] We believe that the US data have some important advantages. First, the data are produced under common definitions and reporting standards. Second, differences in regulatory regimes across the US are relatively small, although they exist, and their effect on the data is weak relative to international comparisons.[6]

---

[3]From which we borrow some elements in the present analysis.
[4] This line of research is part of the growing literature on the macroeconomic effects of credit market imperfections. See Gertler (1988) for a survey of this area.
[5]An exception to this proposition is the work by Faini et all (1993) which look at regions within Italy.
[6]Indeed, Sussman (1993) reports some striking differences in interest rates differentials among developed (OECD) countries and differences in regulatory regimes are the most likely explanation.

The paper is organized as follows. Section 2 describes the theoretical framework and Section 3 contains the main theoretical results. Section 4 describes the basic empirical analysis and Section 5 describes the main results and discusses them. Section 6 stresses some of the economic implications of the empirical findings. Section 7 summarizes the paper and the appendices include proofs and data.

2. The Theoretical Framework

2.1. The Model

Consider a small, open, overlapping generations economy with only one good, which is used both for consumption and for investment. There are L individuals in each generation, who live two periods each. In first period of life they work and some also become entrepreneurs. In second period of life they consume. We assume that individuals are risk-neutral and their utility is:

(1) $$U = c - \delta E,$$

where c is second period consumption, E is disutility of the effort of starting a project and δ is 1 if the individual is an entrepreneur and 0 if not.

Production is done in projects (firms), which are operated by entrepreneurs. Each project requires one unit of capital - which fully depreciates after one period - and labor. Projects differ by two characteristics: location and productivity, which are both random. There is a continuum of projects which are uniformly distributed on a circle, with a circumference which equals L. Note that location may represent other than geographical characteristics: product differentiation for example. As for productivity, we assume for simplicity a two state distribution, namely a project can either succeed or fail. In case of success it produces an amount:

(2) $$A\ g(l)$$

of the physical good, where l is the amount of labor, g is a standard production function and A is a productivity coefficient. If the project fails, productivity is zero. The probability of success is p and this event is independent of project location. A is an economy-wide productivity coefficient, but it might differ across economies. Later on in the paper it serves as a main source of output differences between economies.

Let us now describe the exact timing and information assumptions of the model. After working and earning wages, the young individual has to decide whether to invest the required effort and become an entrepreneur. The specific characteristics of the project (location and productivity) are not known at this stage: the entrepreneur is randomly assigned to a project only after effort has been made. Once location is realized, it is common knowledge to all, especially to the providers of external finance. Next, investment in real capital is made. All this occurs in first period of life. In the second period of life productivity is realized, the entrepreneur hires workers and produces. Once productivity is realized, it is private information to the entrepreneur. Monitoring the productivity of a project from the outside is costly. The reader can recognize here a much simpler version of Townsend's (1979) model of costly state verification. It is efficient (see bellow) to delegate the task of monitoring projects to special institutions which we call banks (like in Diamond, 1984). The monitoring technology requires $a(z)$ workers to monitor a project, where z is the distance along the circle from the project to the bank, and $a'(z) \geq 0$ for all z. This assumption means that banks can reduce the cost of monitoring by specializing on a smaller market segment. Banks are set-up at the period when loans are contracted. Set-up costs are B, and this capital fully depreciates in one period.

Banks raise finance by means of one period deposits, bearing an interest rate of r_t. The market for deposits is assumed to be perfectly competitive. We shall consider two possible regimes. One is free capital mobility, where depositors have free access to the world's market. We assume that the world's deposit interest rate is constant through time and equals r. The other regime is of no capital mobility at all, where the domestic deposit market is isolated from the rest of the world, and it is a case similar to a closed economy. In the paper we concentrate on the free capital mobility case, as it seems to us to be more relevant and it suits better our data set. The analysis of the closed economy is added mostly for analytical completeness, and is left to appendix A.

2.2 Financial Intermediation

As described above, each entrepreneur needs 1 unit of capital for his (her) project. Even if the entrepreneur uses his own savings, he still has to raise outside finance, since we assume that the production function g is such, that equilibrium wages are much smaller than 1. For the sake of simplicity we assume, that an entrepreneur finances his project with external finance only, and puts his savings into bank deposits.[7]

The financial structure of our economy emerges endogenously. Observe, first, the bank-entrepreneur contract. Obviously, the bank has to monitor the project sometimes, otherwise the entrepreneur will always default. But the bank should economize on this costly activity. Townsend (1979) has shown - for a much more general case - that the optimal contract is the standard debt contract, where the bank monitors the project only in the default states, and charges a fixed repayment I in all other (solvency) states. Under such a contract it is incentive compatible for the entrepreneur to reveal truthfully the realized project productivity. Note, that the bank monitors entrepreneurs who claim bankruptcy, although it knows that they are telling the truth, in order to deter from

[7]This assumption does not alter the equilibrium of the economy at all, as can be easily verified. It simplifies the calculations of a measure of the cost of intermediation (which is defined bellow).

cheating. Hence, banks ability to pre-commit must be assumed. Random monitoring strategies by the bank are ruled out as well.[8]

Financial intermediation emerges endogenously in this model as well. The information produced by monitoring a project is valuable to all the financiers of a project. Since monitoring is costly, this creates an incentive for all financiers of the project to delegate monitoring to a single institution, a financial intermediary, which we call a bank (like Diamond, 1984). Moreover, by pooling together many independent projects, the delegated monitor avoids any agency problems *vis a vis* his own external financiers, i.e. the depositors.

2.3. Equilibrium Debt Contracts

Note, that in equilibrium there is a finite number of banks, N_t, in each period t, due to the lump setup costs per bank. It is assumed that they locate on the circle in equal distances from one another. Let $2d_t$ be the distance between any two neighboring banks, then: $2d_t = L/N_t$.

The amount of repayment I depends on the distance between the bank and the project, as is seen in figure 1, which shows how much labor two neighboring banks need in order to monitor a defaulting project - as a function of the location of that project. It is evident that each bank has a cost advantage in monitoring any project within a distance of d_t. Within that domain, a bank can undercut any offer made by any other bank, and monopolize that market segment. We further assume that the bank has a strong bargaining position over the entrepreneur, and can extract all the gains from trade[9]. Hence, the bank acts as a perfect price discriminator and charges the minimum amount, which the closest competitor can charge. Hence,

[8]Random monitoring is usually ignored because it generally yields an optimal contract which does not resemble the standard debt contract (see Mookherjee and Png, 1989). This problem does not exist when the monitored variable has only two realized states of nature, like in our model. Since we hope to capture the more general case, we also ignore random monitoring.

[9]It can be shown (see Sussman 1993) that this is a Nash equilibrium in a game where both banks make simultaneously a take it or leave it offer to the firm.

the equilibrium repayment in period t, for a project at distance z from the bank, $0 \leq z \leq d_t$, $I_t(z)$, equals the rate which equates the marginal profit of the neighboring bank to zero, namely:

$$(3) \qquad I_t(z) = \frac{R_t}{p} + \frac{1-p}{p} w_{t+1} \, a(2d_t - z),$$

where w_{t+1} are expected wages in period $t+1$ and R_t is gross deposit interest rate in period t: $R_t = 1+r_t$. Note, that a bank extracts a high rent from entrepreneurs located next to itself, but breaks even on entrepreneurs located on the margins of the market segment.

2.4. The Equilibrium Conditions

Knowing banks' revenues and costs, we can calculate banks' expected profits, and find the entry condition for banks, at zero profits. Let K_t be the amount of capital invested in production (excluding banking), which is also the number of projects initiated in period t. Let $k_t = K_t/L$ denote the density of projects on the circle, which is also the capital-labor ratio in production. Banks enter until their expected profits in the second period, which are repayments I from solvent projects minus monitoring costs of defaulting projects, equal the gross alternative returns on the setup costs $B\ R_t$. According to (3) we get:

$$(4) \qquad B\ R_t = (1-p)\ k_t\ w_{t+1}\ \phi(d_t) \ ,$$

where ϕ is the following function of half the distance between banks d:

$$(5) \qquad \phi(d) = 2 \int_0^d [a(2d-z) - a(z)] \mathrm{d}z \ ,$$

and is equal to the shaded area in figure 1. Note, that ϕ is increasing and hence condition (4) determines the distance d_t uniquely, and thus determines the number of banks N_t..

Each successful firm produces an amount A g(l) so that profit maximization leads to:

$$(6) \qquad A\ g'(l) = w \ ,$$

where w is wage rate at the time. This condition determines the demand for labor l^d per project, which is decreasing with w/A. The overall demand for labor in the production sector in period $t+1$ is $p\ k_t\ l^d(w_{t+1}/A)$. The demand for labor by banks, for monitoring, is $(1\text{-}p)\ N_t\ k_t\ 2 \int_0^{d_t} a(z)\mathrm{d}z$. The two sectors add up to the overall demand for labor, which equals - in equilibrium - the supply of labor, L. The equilibrium condition in the labor market, at $t+1$, in per capita terms is, therefore:

$$(7) \qquad p\ k_t\ l^d\left(\frac{w_{t+1}}{A}\right) + (1\text{-}p)\ k_t\ \theta(d_t) = 1 \ ,$$

where:

$$(8) \qquad \theta(d) = \frac{1}{d_t} \int_0^{d} a(z)\mathrm{d}z \ .$$

We next derive the entry condition of entrepreneurs. Remember that they decide on entry before location or productivity are realized. Thus, entrepreneurs enter as long as their next period net expected profit - unconditional of location or productivity - but given the perfectly forecasted next period wage rate - exceeds the disutility from effort E. Hence the entry condition in period t is:

$$(9) \qquad p\ A\ \pi\left(\frac{w_{t+1}}{A}\right) - (1\text{-}p)\ w_{t+1}\ \psi(d_t) - R_t - E = 0 \ ,$$

where π is the following profit function:

$$(10) \qquad \pi\left(\frac{w}{A}\right) = g\left[l^d\left(\frac{w}{A}\right)\right] - \frac{w}{A}\ l^d\left(\frac{w}{A}\right) \ ,$$

and ψ is given by:

$$(11) \qquad \psi(d) = \frac{1}{d} \int_0^{d} a(2d\text{-}z)\mathrm{d}z \ .$$

Note, that the first term in equation (9) is expected profits from the project. The second and third terms are (respectively) the expected costs of monitoring (see bellow), and the gross deposit rate, and they add-up to the expected cost of finance.

We now turn to describe the equilibrium in the deposit market. In the case of free capital mobility, the domestic deposit interest rate equals the world interest rate r, due to arbitrage considerations. Hence, the gross deposit interest rate is:

$$R_t = R = 1 + r \,. \tag{12}$$

This is the case we analyze in the rest of the theoretical part of the paper. As for the case of no capital mobility, which is equivalent to a closed economy, saving must equal investment in production and in the financial sector. Hence:

$$w_t = k_t + \frac{B}{2d_t} \,. \tag{13}$$

The rest of the analysis of this case is presented in appendix A.

2.5. The Cost of Financial Intermediation

The above equilibrium conditions enable us to characterize the general equilibrium of the economy in any regime. But before, we introduce another variable, the cost of financial intermediation.

Definition: The cost of financial intermediation in period t, which is denoted by f_t, is the total costs of the bank, labor and capital costs, per one unit of lending to projects by the bank.

Due to zero profits of banks, their costs equal their aggregate revenues from successful projects. It can, therefore, be shown that the cost of financial intermediation in period $t+1$ equals:

$$f_{t+1} = (1\text{-}p)\ w_{t+1}\ \psi(d_t), \tag{14}$$

where ψ is defined by (11). Note, that by use of this variable we can rewrite the entry condition for projects (9), to be:

$$p\ A\ \pi(\frac{w_{t+1}}{A}) - f_{t+1} - R_t - E = 0 \,. \tag{15}$$

Equations (14) and (15) describe the basic feedback effects in our theory. Capital is affected by the cost of intermediation f, as shown in (15), since lower costs raise profitability and induce capital accumulation. The opposite effect, of capital on the cost of financial intermediation, which is described in (14), is less clear. On the one hand, intermediation is labor intensive and hence its cost depends positively on the wage rate w_{t+1}. This is the wage effect. On the other hand, real growth raises bank profits and induces entry of banks. That reduces distance d, as shown in (4), and thus lowers f. This is the banks' specialization effect. These two effects, therefore, work in opposite directions. In the next section we examine what determines their relative effectiveness.

3. The Theoretical Results

We first verify the existence and stability of the equilibrium in the economy with full capital mobility.

Proposition 1: The economy reaches the steady state instantaneously. If the banking sector is sufficiently small, namely if B and 1-p are sufficiently small, the equilibrium is unique and stable.
The proof is in appendix B.

Since the economy is always at the steady state, we delete from now on all time subscripts, as all variables are constant through time. The endogenous equilibrium variables in this economy are capital per capita k, the wage rate w, half the distance between banks d, and the cost of financial intermediation f. The exogenous variables are the economy specific productivity coefficient A, the probability of success p and the gross interest rate on deposits R. Our main goal is to find how the equilibrium depends on these exogenous variables. We concentrate in the analysis mainly on the effect of A, since we view it as the main source of differences in output between economies. Hence, changes in A capture the effect of growth, especially in a cross-section comparison of economies.

In order to solve the model, we reduce the equilibrium conditions to a set of two equations in the two variables w and f, by simple manipulations. The first is derived from the zero profit condition for projects, equation (15):

$$(16) \qquad f = p\, A\, \pi(w/A) - R - E\ ,$$

and the second is derived from the labor market equilibrium condition:

$$(17) \qquad p\, l^d(w/A) + (1\text{-}p)\, \theta\{\psi^{-1}[\frac{f}{(1\text{-}p)w}]\} = \frac{(1\text{-}p)w}{BR}\, \phi\{\psi^{-1}[\frac{f}{(1\text{-}p)w}]\}\ .$$

These two equations together determine uniquely the equilibrium values of w and f, and enable us to examine the effect of changes in the exogenous variables. As already explained, such comparative statics must yield ambiguous results, due to the opposing effects of wage costs and of specialization. In order to assess more precisely the relative weights of the two effects we use a specific parameterization of technology in the real and the financial sectors. Let us assume that real production is described by a Cobb-Douglas production function, namely that g is:

$$(18) \qquad g(l) = \frac{1}{\alpha}\, l^{\alpha}\ .$$

Let us further assume that monitoring labor requirements are:

$$(19) \qquad a(z) = \frac{a}{\beta+1}\, z^{\beta}\ ,$$

for all z, where $\beta \geq 0$. The β parameter is a measure of the curvature of a, and it measures the returns to specialization. In the $\beta = 0$ case we assume that capital requirements B are 0, to rule out the case of a natural monopoly.

The main claim of the paper, of feedback between the real and the financial sectors, due to specialization, can be phrased in terms of the the effect of productivity on growth. The financial feedback magnifies the effect of changes in A on output as shown in Proposition 2.

Proposition 2: The effect of A on w is positive and it increases with β , namely $\frac{\partial w}{\partial A}$ increases with β.
The proof is in appendix B.

Although Proposition 2 states the main theoretical result of the FDC, it is not phrased in terms of observable variables. This can be done by concentrating on the cost of financial intermediation f, both because it is observable, and because it is the channel through which the feedback effect operates. Indeed, Proposition 3 shows that the greater the returns to specialization, the stronger is the specialization effect of A on f.

Proposition 3: The specialization effect of A on f rises with β, namely $\frac{\partial f}{\partial A}$ declines with β. More specifically:

(a) If $\beta < 1$ the wage effect dominates and f rises with A.

(b) If $\beta = 1$ the two effects cancel one another and f is independent of A.

(c) If $\beta > 1$ the specialization effect dominates and f declines with A.

The proof is in appendix B.

From propositions 2 and 3 together we learn that the effectiveness of specialization can be assessed by measuring the correlation between output, or wages w, and the cost of financial intermediation f. This is indeed the main empirical test we run, in section 4. But before we turn to that, let us examine the predictions of the model on the effects of the other variables on f.

We first examine the effect of R on f. Note, that in a world with free and frictionless capital mobility all economies face the same deposit interest rate R. But in reality this is only a first order approximation. There are three separate effects of R on f. The first is the wage effect, since a rise in R reduces capital and wages, and thus reduces f. The second is the specialization effect, as a higher R and less capital drive banks out, and that raises f. The third effect is a direct cost of capital effect, due to banks' setup costs $B\ R$. Proposition 4 shows that the overall effect depends, again, on the curvature of the a function. For the sake of simplicity we concentrate on two opposite cases of monitoring costs, the fixed and the linear, namely $\beta = 0$ and $\beta = 1$, respectively.

Proposition 4: A rise in R reduces f if monitoring costs are fixed, and it raises f if monitoring costs are linear.
The Proof is in appendix B.

The effect of the probability of default $1\text{-}p$, on the cost of financial intermediation $f = (1\text{-}p)\ w\ \psi(d)$, also comprises of three separate effects. The first is the direct effect, as more defaults call for more monitoring, which increases f. The second is the wage effect, which operates in the opposite direction, since a higher default rate $1\text{-}p$ decreases the demand for labor and reduces wages. The third effect is specialization, as exit of banks reduces specialization and raises f.

Proposition 5: A rise in $1\text{-}p$ lowers f in the case of fixed monitoring costs, as long as $1\text{-}p \leq \alpha$. In the case of linear monitoring costs a rise in $1\text{-}p$ raise f.
The proof is in appendix B.

We can therefore summarize the theoretical analysis in the following set of equations, which describe the reduced form of the system:

(20) $$w = W(A,\ 1\text{-}p,\ R)\ ,$$

and:

(21) $$f = F(A,\ 1\text{-}p,\ R)\ .$$

The partial derivatives of both W and F depend on the returns to specialization, on the strength of the specialization effect. Hence, any empirical evidence on these derivatives, especially those of F, can teach us on the effectiveness of the specialization effect and, therefore, on the FDC.

4. The Empirical Analysis

The objective of the empirical analysis is to estimate a linear approximation of the reduced form equation (21), from a panel of US banking data consolidated at a state level. Because of data limitations, some adjustments have to be made. So practically, we estimate the following equation:

$$(22) \quad COST/GLL = \alpha_1 + \alpha_2\ WRTOFF/GLL + \alpha_3\ DEPO/GLL + \alpha_4\ Log(GSP/CAPIA) + \alpha_5\ X + \varepsilon,$$

where COST is bank's operational expenses, GLL outstanding credit, WRTOFF the amount of bad debt written off by banks, DEPO non interest bearing deposits, GSP is gross state product (the state equivalent of GDP)[10], CAPITA the state's population[11], X a vector of some possibly important exogenous variables which are absent from the theoretical analysis, and ε an error term. The source of all banking data is the Federal Deposit Insurance Corporation (FDIC)[12], which publishes (state consolidated) balance sheets and income statements of banks under its supervision.

As noted above, some adjustments had to be made in equation (22) to overcome some data limitations. The first of these limitations has to do with the fact that

---

[10]Source: Rensaaw et. al. (1988) and Trott et. al. (1991).
[11]Source for population: Statistical Abstracts of the United States.
[12]Statistics On Banking an annual publication) and Historical Statistics on Banking.

there are no observations on A, the state productivity level which is responsible to the real differences among states. We approximate A by GSP/CAPITA. This approximation is problematic because according to our theory state income is affected by financial factors which may cause endogeneity problems. Ideally we should have estimated simultaneously an income equation and a cost of intermediation equation. Unfortunately, we could not find satisfactory instruments for that purpose.

The COST variable includes all operational expenses on workers, buildings, hardware etc.[13]. It should be stressed that COST does not include interest expenses. Divided by extended credit - or "gross loan and leases" (GLL) in the FDIC's terminology - we get a variable (COST/GLL)[14] which can be defined as the cost of banking: total banking cost per $1 intermediated between an ultimate lender and an ultimate borrower. The cost of banking should not be confused with the cost of financial intermediation - f of our theoretical model. The reason is that banks are multi product firms. They produce (at least) two services: intermediation and clearing services[15]. As we cannot observe the cost of intermediation separately, we add the variable DEPO/GLL into equation (22), so that it would serve as a proxy for the volume of clearing services produced.

WRTOFF is the amount of bad debt which is written off the banks' balance sheet within a year - "charge-offs" - in the FDIC's terminology. Divided by GLL it gives us the *ex post* probability of default, or $1-p$ in terms of our theoretical

---

[13]Including expenses on labor, fixed capital and "other expenses." Other expenses include directors' fees, legal fees, insurance premiums, advertisement and public relations, and also some capital loses due to the sail of loans and future or forward contracts (see p. 73 in the Historical Statistics). Ideally, some of these expenses should be excluded, but the relevant decomposition is not published.
[14]Note that the numerator is a flow variable, while the denominator is a stock variable. All variables of this sort where adjusted to inflation under the assumption that the flow is uniformly distributed along year. Specifically, it was multiplied by $[(1+\pi)\cdot \mathrm{Log}(1+\pi)]/\pi$ where π is yearly rate of inflation.
[15]Indeed, the question of how to measure the bank's output, is one of the central issues dealt by the empirical work in banking. See Gilbert (1984), Berger, Hanweck and Humphrey (1987), Clark (1988) and especially Kim (1986).

model. The deposit interest rate R was eliminated from equation (22) because according to our model all states should have the same deposit rate.[16]

The data appear under common formats and definitions, and seem to be relatively consistent for the years 1984-1991. This is a problematic sample period. First, it does not cover even a single business cycle. Second, it is a period of unprecedented turmoil in the post (second world) war history of the US banking industry: deregulation, stock market instability, a wave of financial innovations and the S&L debacle. We shall thus give an extra attention in the analysis to the year 1985. It is a mid cycle year, well into the economic recovery of the 1980's, but still preceding the slump of the end of the decade. 1985 follows the period of fast deregulation in the beginning of the decade, but it precedes the 1987 crash and the financial instability that followed. We suggest, however, that results obtained from such a sample period should be considered cautiously.

Table 1 provides some descriptive statistics of 1985's data. It is immediately visible that the cost of banking is strikingly high: ¢5.5 on a dollar of extended credit. In states like Colorado or Wyoming the cost of banking exceeds ¢7 on the dollar (see Appendix C). The magnitude of these numbers can be better appreciated against a flow variable. In 1985 banks have earned on average ¢15.8 in interest and fees on a dollar of extended credit. Thus, about a third of banks' income is being eaten up by their operational expenses[17]. Table 1 provides information about some additional magnitudes. Of special interest is the realized probability of default: 1.15% on average for 1985.

[16]To the best of our knowledge there is no data on states' interest rates.
[17]Most of the rest is paid as interest on deposits.

Table 1: Descriptive Statistics, 1985

	Average	S.D.
GSP/CAPITA	$16,175	$2478
COST/GLL	5.4%	0.7%
Interest & Fees/GLL	15.8%	1.9%
Interest payments/GLL	9.4%	1.4%
DEPO/GLL	29.7%	6.7%
WRTOFF/GLL	1.15%	0.8%
Interest rate differential (prime lending)	1.5%	-

More detail about the regional distribution of the cost of banking can be found in the map presented in Figure 3. The regional pattern seems to be quite strong: the cost of banking is low in the north east and high in the west. This may be a result of differences in the level of real economic development in the various regions, differences in sectorial structure (farming, manufacturing etc.), or pure spatial factors like distances from financial centers (see bellow).

Our sample includes all mainland US (48) states. Hawaii, Alaska and DC are excluded by that definition. Since all of them have some outlying characteristics[18], their inclusion is probably undesirable. We did not, however, exclude other outlyers, which are covered by the above definition. Most noteworthy are Delaware and South Dakota, who have a very low DEPO/GLL ratio, as these states host some functions of out of state banks: they are exporters of banking services. Excluding them did not change the results in any dramatic way.

[18]Alaska and DC have an exceptionally high GSP/CAPITA.

5. Estimation

Table 2 (line 1) presents an estimation of equation (22) (numbers in brackets are t statistics based on consistently estimated White standard errors; N is the number of observations). The estimation is consistent with the theoretical model and gives support to the FDC, since it shows that a specialization effect not only exists, but it is strong enough to overcome the direct wage cost effect.

Table 2: The Cost of Banking, 1985

Dependent Variable: Cost/GLL

	Method	Intercept	WRTOFF/GLL	DEPO/GLL	Log(GSP/CAPITA)	R^2	N
1	OLS	6.64 (3.42)	0.51 (5.19)	0.07 (5.95)	-1.36 (-1.90)	0.57	48
2*	TSLS	6.64 (4.58)	0.52 (4.51)	0.07 (6.13)	-1.36 (-2.53)	0.53	48

*In this estimation, regular rather than White t statistics are presented.

Line 2 handles a possible endogeneity problem in WRTOFF/GLL. Suppose - realistically - that project income has a continuous distribution rather than a two state distribution, as in our simple theoretical model. In that case the probability of default is endogenously determined: the higher the interest rate, the higher is the probability of default. To handle this problem we need some instruments, which affect the whole distribution function of project income, and affect the probability of default independently of the other endogenous variables. It is likely that this distribution is affected by the sectorial structure of the state economy, and the phase of the business cycle at which the state economy is currently in. We therefore use the share of farming and manufacturing in GSP[19] and the rate of change of GSP (FARM and MANUF and GRATE respectively) as instruments. The estimation is only marginally affected. Indeed, when we estimated the default

[19]Source: Rensaaw et. al. (1988) and Trott et. al. (1991).

equation, COST/GLL came out insignificantly and had a wrong sign. Apart from that the equation looks quite good.

$$\underset{(3.22)}{WRTOFF/GLL = 11.92} - \underset{(-1.01)}{0.18} (COST/GLL) + \underset{(4.02)}{0.11} FARM - \underset{(-3.04)}{0.04} MANUF - \underset{(-2.86)}{8.41} GRATE.$$

We can now turn to some other possible effects which are not dealt with explicitly in our simplifying theoretical model. The regional pattern which is visible in Figure 3 can be interpreted - in the spirit of our theoretical model - as reflecting distances from financial centers. We define the variable DIST≡Min(distance from NY, distance from LA, distance from Chicago)[20]. Line 1 in Table 3 gives some support to this view as the distance tends to raise cost of banking. Alternatively, we define a variable GLL/GSP, which is supposed to be high in financial centers, because some of the lending is directed to out of state customers. This variable also comes out very significantly (line 2) and seems to indicate a lower cost of banking at financial centers.

Table 3: Additional Variables, 1985

OLS, Dependent Variable :COST/GLL

	Intercept	WRTOFF/ GLL	DEPO/ GLL	Log (GSP/ CAPITA)	Log(DIST)	GLL/GSP	MANUF	R^2	N
1	4.98 (2.21)	0.45 (5.30)	0.07 (5.45)	-0.90 (-1.15)	0.09 (2.12)			0.60	48
2	6.91 (4.37)	0.52 (5.73)	0.04 (2.72)	-1.01 (-1.56)		-9.35 E-4 (-3.87)		0.64	48
3	8.46 (4.89)	0.36 (4.08)	0.07 (7.61)	-1.69 (-2.66)			-0.03 (-4.14)	0.64	48
4	8.31 (5.37)	0.39 (4.68)	0.05 (3.18)	-1.34 (-2.16)		-7.45 E-4 (-2.93)	-0.03 (-3.72)	0.68	48

[20]The distance is measured to the state capital or major city.

We also test whether some sectors are more costly to monitor than others by adding indicators of the sectorial structure into the regression. Line 3 gives some support to the idea that manufacturing is cheaper to monitor than other sectors of the economy. When we add all three variables to the regression, the DIST variable comes out insignificant. Hence, line 4 seems to be our best estimation for the cost of banking. We call it henceforth the 'basic equation'.

We now pool the cross sections of 1984-1991 and estimate our basic equation from the panel derived. The advantage of such a procedure is not just a larger number of observations, but that adding state (and time) dummies eliminate part of the potential bias from specific state characteristics, which are not captured by the variables discussed in Table 3. Our results are presented in Table 4.

Table 4: The Cost of Banking, 1984-1991

OLS, Dependent Variable :COST/GLL, with state and time dummies

Intercept	WRTOFF/ GLL	DEPO/ GLL	Log (GSP/ CAPITA)	GLL/GSP	MANUF	R^2	N
12.11 (3.99)	0.50 (8.79)	0.13 (10.27)	-3.47 (-3.90)	-6.06 E-5 (-0.16)	-.05 (2.91)	0.83	384

Overall, the results are quite similar to those obtained with 1985 data. The main significant difference is that DEPO/GLL has a coefficient which is almost double the one estimated from 1985 data. At least some of the explanation may be due to the fact that DEPO/GLL is decreasing in the sample period from 30% in 1984 to 23% in 1991. It is unlikely that the volume of clearing services has changed significantly in that period. Thus, each dollar of DEPO represents more clearing services, and thus induces greater cost of banking.

In a sense, the results presented in Table 4, tend to create an impression of over robustness. When we estimate the same equation separately for each year, we find some substantial differences across years. These differences are averaged out when the yearly observations are pooled together. They happen to resemble the 1985 based estimation.

6. An Economic Interpretation of The Results

6.1 The Cost of Financial Intermediation

We believe that the results of section 5 provide strong evidence in favor of the Diamond-Townsend theory: the amount of bad debt has a strong (positive) and highly significant effect on the cost of banking. Moreover, the coefficient of WRTOFF/GLL is an aproximation of the cost of monitoring[21]. The magnitude of this cost is striking: most of our estimates lie within a confidence interval around ¢40 on a dollar of bad debt.

This number is higher by an order of magnitude than the estimates of bankruptcy costs, which are usually cited in the corporate finance literature. Warner (1977), in his well known study of direct bankruptcy costs in failed railroad companies, argues that 3% (of market value prior to default) would "overstate the ... apparent direct cost of bankruptcy." This statement reflects a wide belief in the corporate finance literature that bankruptcy costs are small and probably insignificant. Two differences between these studies and ours should be stressed. First, we measure the cost, which bankruptcy imposes on the lender, while the corporate finance literature usually measures the cost imposed on the defaulting corporation. Second, the corporate finance literature (c.f. Warner) often focuses on large publicly held corporations. It is quite likely that low bankruptcy cost is one of the very factors which allow those corporations to go public in the first place.

---

[21]Differentiating (14) by $(1\text{-}p)$ (omitting the time indices) we get $df/d(1\text{-}p)=w\psi+(1\text{-}p)[\psi dw/d(1\text{-}p)+w\psi' dd/d(1\text{-}p)]$. $w\psi$ is the average cost of monitoring. The other elements are multiplied by $(1\text{-}p)$ and are likely to be very small.

The cost of monitoring, as high as it is, still explains a relatively small part of the cost of banking. Consider again the mid cycle year - 1985. Since banks wrote off (on average in 1985) just about 1.15% of their extended credit (GLL), bankruptcy costs can explain a little less than ¢0.5 of the ¢5.4 on the dollar (i.e about 9%) of the gross cost of banking. On the other hand, substituting the average level of DEPO/GLL reported in Table 1 into the second estimation of Table 2, shows that clearing services are responsible for only about ¢2 (on a dollar of extended credit) of the gross cost of banking. This leaves a substantial part of the cost of banking which cannot be related to either the cost of clearing services or *ex post* monitoring. It is possible, however, that there are some additional costs of intermediation beside monitoring: evaluating prospects *ex ante*, processing applications, supervising projects on an on-going basis, marketing deposits etc. To partially answer this question, we run our basic equation in absolute values, without deflating them by extended credit. The results appear in Table 6.

Table 5: The Basic Equation in Levels, 1985

OLS, Dependant Variable: COST

Intercept	WRTOFF	DEPO	GLL	Log(GSP/CAPITA)	GLL/GSP	MANUF	R^2	N
578128 (1.02)	0.70 (1.86)	0.08 (2.27)	0.017 (1.86)	-187616 (-0.97)	-94.73 (-0.76)	4018 (1.03)	0.99	48

The estimation seems to be somewhat problematic: GLL/GSP and MANUF are not significant, and the coefficient of WRTOFF is very high. Note also that the estimation is affected by the fact that the data are not descaled. This immediately blows up the variance of the dependent variable, which is "explained" to a great extent by the scale of the state; hence the high R^2. The other costs of intermediation are 1.7¢ on a dollar intermediated but the coefficient is hardly significant.

Does this number look realistic? to support it, observe that the interest rate differential on prime lending (Table 1)[22], was 1.5% in 1985. Since this is an interest rate differential on safe lending, the default premium and the expected cost of monitoring should be insignificant. The differential here, therefore, reflects those other costs of intermediation. The number looks surprisingly close to the one estimated in Table 5.

We can thus conclude that of the cost of banking (¢5.5 on the dollar), about ¢2 can be attributed to clearing services, ¢2 to intermediation, and the rest are either fixed costs (common to both clearing and intermediation), or can be attributed to one of the two activities only by a more refined econometric technique (and better data).

6.2 Banking and Development

The estimations in section 5 have already established a significant relationship between real economic development and the cost of banking. The coefficient is around -1. Since GSP/CAPITA in 1985 varies from around $12,000 per annum in poor states (like Mississippi or Arkansas) to around $19,000 per annum in rich states (like New York or California), our conclusion is that gross cost of banking should be lower by about ¢0.45 on a dollar in the rich states relative to the poor states.

It is worth stressing that the coefficient of Log(GSP/CAPITA) is not a full measure of financial specialization, but rather a measure of how much its effect dominates the direct wage effect. Hence the specialization effect should even be stronger.

Remember also that in reality, there are other characteristics of the real economy which may affect the cost of banking. One example is the degree of

---

[22]This is $(1+i_L)/(1+i_D)-1$ where i_L i_D are the prime rate and the 3 month Euro rate respectively. Note that this calculated this way, inflationary expectations cancel. Source: Federal Reserve Bulletin.

competition among banks. To the extent that competitiveness is determined by the development level, as in our model, this effect is already captured in the variable GSP/CAPITA. But we can consider the additional effect of state specific level of competition among banks. We capture this by a variable of CNTR, which measures concentration of the banking sector in a state. Since the FDIC decomposes the consolidated balance sheets of state's banking industry by size groups of banks, we can calculate a Hirfindahl index under the assumption that all banks in each size group are of the same size[23]. It can be shown (line 1 Table 7) that the cost of banking tends to be high in states where the banking system is relatively concentrated.

Table 6: Additional Factors, 1985

OLS, Dependent Variable :COST/GLL

	Inter-cept	WRTOFF/GLL	DEPO/GLL	Log(GSP/CAPITA)	GLL/GSP	MANUF	CNTR	Log(BRANCH)	R^2	N
1	7.48 (5.18)	0.44 (5.31)	0.05 (3.46)	-1.20 (-2.08)	-7.21 E-4 (-2.81)	-0.02 (-3.07)	1.64 (2.21)		0.70	48
2	8.00 (5.56)	0.56 (4.79)	0.05 (3.18)	-1.42 (-2.36)	-8.05 E-4 (-2.83)	-0.03 (-3.20)		0.18 (1.92)	0.71	48

The issue of branching is not directly related to our analysis. Still, for the sake of completeness, and to check that the omission of that factor does not bias our results, we add the average number of offices (including head office) per bank (BRNCH). Our results do not support the claim that restrictions on branching impose a higher cost of banking.

[23]More accurately, if a_i is the amount of extended credit (GLL) provided by all the banks of size group i, and n_i the number of banks within size group i, $CNTR=\Sigma_i n_i[a_i/n_i)/A]^2$, where $A=\Sigma_i a_i$.

6.3 Some Business Cycle Effects

When the US economy slumped into recession by the end of the decade, a state of financial distress have developed where both the cost of banking and the amount of bad debt increased substantially. Nowhere has the effect of the recession on the banking system been more dramatic than in Texas (see Figure 5). It had one of the nation's lowest default rates and cost of banking in 1984-5; but WRTOFF/GLL peaked 3.1% in 1988, and COST/GLL reached 8.4% in 1991. The same trend can be observed on the national cross state averages, even though the magnitudes are not that dramatic.

The effect of entering the recession on the cost of banking can also be seen in figure 6, which plots the coefficients of the time dummies from the panel regression in Table 4. The results speak for themselves. The implications of our theory to business cycle analysis look promising, but are left to future work. It also serves as a reminder to the fact that our analysis ignores some important factors, a potential source for misspecification and biases in the analysis.

6. Summary and Conclusions

This paper supplies, we believe, strong evidence in favor of the financial development hypothesis, namely that there are strong feedback effects between the real and financial sectors.

Our analysis of the relationship between the two sectors is based on two main assumptions, that financial intermediation is a costly activity, and that these costs decline with specialization. Thus, economic development, which raises banks' profits and induces more entry of banks, increases specialization and reduces the costs of intermediation. This is the mechanism through which the feedback effect works.

We empirically examine the feedback effect by measuring the correlation between the real sector and the cost of financial intermediation. The data we use, of cross state banking sectors in the US, shows that the cost of financial intermediation is negatively related to real activity. That means that the specialization effect is strong and so is the feedback effect.

References

Atje, R. and Boyan Jovanovic (1992). "Stock Markets and Development." mimeo.

Berger, Allen N., Gerald A. Hanweck and david B. Humphrey (1987). "Competitive Viability in Banking," Journal of Monetary Economics, (20), pp. 501-520.

Cameron, Rondo, (1967). "Banking and Development in the Early Stages of Industrialization." New York: Oxford University Press.

Clark, Jeffrey A., (1988). "Economics of Scale and Scope at Depository Financial Institutions: A Review of the Literature," Federal Reserve Bank of Kensas City Economic Review, September/October, pp. 16-33.

Diamond, Duglas, (1984). "Financial Intermediation and Delegated Monitoring." Review of Economic Studies 51, pp. 393-414.33

De la Fuente, Anjel and Jose Maria Marin (1993). "Innovation, "Bank" Monitoring and Endogenous Financial Development." mimeo.

Faini, Ricardo, Giampaolo Galli and Curzio Giannini (1993). "Finance and Development: The Case of Southern Italy," In Alberto Giovaninin (ed). Finance and Development: Issues and Experience, Cambridge: Cambridge University Press.

Gale, Duglas, and Martin Hellwig (1985). "Incentive-Compatible Debt Contracts: The One Period Problem." Reviw of Economic Studies 52, pp. 647-663.

Gertler, Mark (1988). "Financial Structure and Aggregate Economic Activity: an Overview." Journal of Money Credit and Banking 20, pp. 559-588.

Greenwood, Jeremy and Boyan Jovanovic (1990). "Financial Development, Growth, and the Distribution of Income." Journal of Political Economy, 98 pp. 1076-1107.

Gilbert, Alton R. (1984). "Bank Market Structure and Competition," Journal of Money Credit and Banking, 16, pp. 617-645.

Goldsmith, Raymond (1969). Financial Structure and Development. New Haven: Yale University Press.

Gurley, John G. and Edward S, Shaw (1955). "Financial Aspects of Economic Development." American Economic Review 45, pp. 515-538.

Lucas, Robert E. Jr., (1990). "Why Doesn't Capital Flow From Rich to Poor Countries?" American Economic Review, 80.

Mookerherjee, Dilip and Ivan Png (1989). "Optimal Auditing, Insurance, and Redistribution." Quarterly Journal of Economics, May, pp 399-415.

Pagano, Marco (1991). "Financial Markets and Growth: an Overview." Memo.

Kim, Moshe, (1986). "Banking Technology and and Existence of a Consistent Output Aggregate," Journal of Monetary Economics, 18, pp. 181-195.

Kindleberger, Charles P. (1974). "The Formation of Financial Centers: A Study in Comparative Economic History," Princeton Studies in International Finance No. 3.

Kravis, Irving B. and Robert E. Lipsey, (1983). "Towards an Explanation of National Price Levels." Princeton Studies in International Finance, No. 52.

Rensaaw, Vernnon, Edward A. Trott Jr. and Howard L. Friedenberg (1988). "Gross State Product by Industry 1963-86." Survey of Current Business, 68 pp. 30-46.

Salop S. (1979). "Monopolistic Competition with Outside Goods." Bell Journal of Economics, 10, pp. 141-156.

Shaw, Edward (1973. Financial Deepening in Economic Development. Oxford, Oxford University Press.

Sussman, Oren (1993). "A Theory of Financial Development," In Alberto Giovaninni (ed). Finance and Development: Issues and Experience, Cambridge: Cambridge University Press.

Townsend, Robert M. (1979). "Optimal Contracts and Competitive Markets with Costly State Verification." Journal of Economic Theory 21, pp. 265-293).

Trott, Edward A. Jr., Ann E. Dunbar and Howard L. Friedenberg (1991). "Gross State Product By Industry 1977-89." Survey of Current Business, 71, pp. 43-59.

Warner, Jerold B. (1977). "Bankruptcy Costs: Some Evidence." Journal of Finance, Vol. 32.

Zilibotti, Fabrizio (1993). "Endogenous Growth and Intermediation in an Archipalgo Economy." mimeo.

Appendix A

In this appendix we briefly describe the dynamics of the economy with no capital mobility, and examine the conditions under which f rises or falls during the development path. The dynamics of capital accumulation are described by the set of equations (4), (7), (9) and (13). In each period t these equations fully determine the equilibrium variables d_t, k_t, w_{t+1} and R_t, as functions of the predetermined variable w_t. For the sake of simplicity, we assume that the productivity coefficient A is equal to 1.

By some manipulations we can reduce the dynamic system to two equations, with two dynamic variables w_{t+1} and d_t. The labor market equilibrium condition can be written as:

$$\text{(A.1)} \qquad [p\, l^d(w_{t+1}) + (1\text{-}p)\, \theta(d_t)]\, (w_t - \frac{B}{2d_t}) = 1 \, ,$$

and the condition of zero profits for projects becomes:

$$\text{(A.2)} \qquad p\, \pi(w_{t+1}) - \frac{1\text{-}p}{B}\, w_t\, w_{t+1}\, \phi(d_t) - (1\text{-}p)\, \theta(d_t) - E = 0 \, .$$

We now define a steady state in this economy, as a dynamic equilibrium, in which all variables are constant. In terms of the reduced form system, a steady state is a pair (w, d), which satisfies the following conditions:

$$\text{(A.3)} \qquad [p\, l^d(\mathrm{w}) + (1\text{-}p)\, \theta(d)]\, (w - \frac{B}{d}) = 1 \, ,$$

and:

$$\text{(A.4)} \qquad p\, \pi(\mathrm{w}) - \frac{1\text{-}p}{B}\, \mathrm{w}^2\, \phi(d) - (1\text{-}p)\, w\, \theta(d) - E = 0 \, .$$

In order to analyze stability of the steady state, we add two assumptions. First we assume that labor demand is more than unit elastic. This assumption is necessary for stability in this model, even without a financial sector and with no intermediation costs. The second assumption is that the financial sector is small relative to the real economy, namely, that B and $1\text{-}p$ are small.

Proposition A.1: If B and $1\text{-}p$ are sufficiently small and if $w\ l^d(w)$ is decreasing in w, a steady state exists and is stable. As the economy converges to the steady state, w rises and d declines.

Proof: If the assumptions of the proposition hold, the L.H.S. of equation (A.4) is a decreasing function of w, which guarantees existence and uniqueness of the steady state.

Equations (A.1) and (A.2) determine w_{t+1} and d_t uniquely. We can calculate the effect of a change of w_t and show that since $\phi(d)/d$ is increasing and if $B < \frac{w_{t+1}}{w_t\, R_t}$ in the neighborhood of the steady state, then:

$$\frac{\partial\ w_{t+1}}{\partial\ w_t} > 0.$$

A calculation of the effect on d_t shows that it always satisfies:

$$\frac{\partial\ d_t}{\partial\ w_t} < 0.$$

Hence, growth raises wages and specialization of banks.

In order to examine stability, let us define a new variable: $x_{t+1} = \frac{w_{t+1}}{w_t}$. If we substitute it in (A.1) and (A.2) we get:

$$\text{(A.5)} \qquad p\ \pi(w_t x_{t+1}) - \frac{1-p}{B}\ w_t^2\ x_{t+1}\ \phi(d_t) - (1\text{-}p)\ w_t x_{t+1} \theta(d_t) - E = 0\ ,$$

and:

$$\text{(A.6)} \qquad [p\ l^d(w_t x_{t+1}) + (1\text{-}p)\ \theta(d_t)]\ (w_t - \frac{B}{2d_t}) - 1 = 0\ .$$

Due to the assumptions of this proposition the L.H.S. of (A.6) is decreasing in w_t. Hence, it can be shown that:

$$\frac{\partial\ x_{t+1}}{\partial\ w_t} < 0\ .$$

Hence, the steady state is stable. Q.E.D.

Proposition A.1 presents us with both the wage effect and the specialization effect, since growth raises wages and reduces the distance between banks at the same time. In order to assess more precisely the two effects, we use the Cobb-Douglas production function (18) and the monitoring function (19) in the two cases of fixed monitoring costs, $\beta = 0$, and linear monitoring costs, $\beta = 1$.

Proposition A.2: If monitoring costs are fixed f rises as the economy grows. In the case of linear costs f declines with growth.

Proof: In the case of fixed monitoring costs the cost of financial intermediation is: $f_{t+1} = (1-p)\, w_{t+1}\, a_0$ and hence it rises as the economy grows and wages rise, according to proposition A.1.

In the case of the Cobb-Douglas production function we can combine equations (A.1) and (A.2) together to get:

$$\frac{2\, d_t w_{t+1}}{2\, w_t d_r\, B} - \frac{1-p}{\alpha}\, w_{t+1}\, \theta(d_t) - \frac{1-p}{B}\, w_t\, w_{t+1}\, \phi(d_t) - E = 0 \,.$$

If we apply this equation to the case of linear costs, and substitute the variable $x_{t+1} = w_{t+1}/w_t$ we get:

$$\frac{2\, f_{t+1}}{2\, f_{t+1} - 3\, a_0\, (1-p)\, x_{t+1}\, B} - \frac{1}{3\, \alpha\, a_0} - \frac{4\, f_{t+1}^{\,2}}{9\, B\, a_0\, (1-p)\, x_{t+1}} - E = 0 \,.$$

This equation establishes a positive relationship between x_{t+1} and f_{t+1}. According to proposition 1, x declines as the economy grows. Hence, f declines as well in this case. Q.E.D.

Appendix B

Proof of Proposition 1

The dynamic equations in period t are:

(B.1) $(1\text{-}p)\ w_{t+1}\ k_t\ \phi(d_t)\ -\ B\ R\ =\ 0\ ,$

(B.2) $p\ k_t\ l^d(w_{t+1})\ +\ (1\text{-}p)\ k_t\ \theta(d_t)\ =\ 1\ ,$

and:

(B.3) $p\ \pi(w_{t+1})\ -\ (1\text{-}p)\ w_{t+1}\ \psi(d_t)\ \text{-}R\ -\ E\ =\ 0\ .$

Hence, period t variables do not depend on any perdetermined variables and the economy moves instantaneously to the steady state.

Next, note that from equation (B.2) we get that next period wages depend on current capital k and distance d (we omit time notation from now on):

(B.4) $w_{t+1}\ =\ w(\ k_t,\ d_t)\ .$

If we substitute (B.4) in equations (B.1) and (B.3) we get a set of two equations with two variables k and d, which also determine the dynamics of banks and production investment, correspondingly:

(B.5) $(1\text{-}p)\ w(\ k,\ d)\ k\ \phi(d)\ -\ B\ R\ =\ 0\ ,$

and:

(B.6) $p\ \pi[\ w(\ k,\ d)]\ -\ (1\text{-}p)\ w(\ k,\ d)\ \psi(d)\ -\ R\ -\ E\ =\ 0\ .$

We now need to show that these two equations determine a unique and stable equilibrium. For that. we need to show that the determinant of the Jacobian is negative. It can be shown that this holds, if $1\text{-}p$ is sufficiently small. Q.E.D.

Proof of Proposition 2

If production is Cobb-Douglas the equilibrium equations (16) and (17) can be transformed to the following two equations:

(*) $f = p\ \frac{1-\alpha}{\alpha}\ A^{\frac{\alpha}{1-\alpha}}\ w^{\frac{-\alpha}{1-\alpha}}\ -\ R\ -\ E\ ,$

and:

$$\frac{\alpha}{(1-\alpha)\,w}\,(f + R + E)\,(1-p)\,\theta(d) = \frac{(1-p)\,w\,\phi(d)}{B\,R}\,.$$

Note, that the first equation does not depend on the monitoring technology at all. Note too that the second equation does not depend on A at all. If monitoring is described by (19), then the second equation can be written as:

$$(**)\qquad [\tfrac{\alpha}{1-\alpha}\,(f + R + E) + \tfrac{f}{\varepsilon}]\,w^{1/\beta-1} = \lambda\,f^{1+1/\beta}\,,$$

where $\varepsilon = 2^{\beta+1} - 1$, and $\lambda = 2\,(\varepsilon-1)\,B^{-1}\,R^{-1}\,\varepsilon^{-1-1/\beta}\,(1-p)^{-1/\beta} > 0$.

Figure 2 describes the curves in the plain (w, f) which are defined by (*) and by (**), respectively. The slope of (**) can be shown to equal:

$$\frac{-\lambda\,(\beta - 1)\,w^{-1/\beta}}{(\frac{\alpha}{1-\alpha} + \frac{1}{\varepsilon})\,f^{1-1/\beta} + \frac{\alpha\,(1+\beta)}{1-\alpha}\,(R + E)\,f^{2-1/\beta}}\,.$$

It is easy to verify that this slope is decreasing with β. A rise in A shifts the (*) curve, without affecting (**). It is clear that it raises w, and it is also clear from figure 2 that the higher is β the greater the effect on w. Q.E.D.

Proof of Proposition 3

Remember from the proof of proposition 2, that as A rises the (*) curve in figure 2 shifts to the right. The effect on f clearly depends on the slope of the (**) curve. It is easy to see from figure 2 that f rises if $\beta < 1$, declines if $\beta > 1$, and is horizontal if $\beta = 1$. Q.E.D.

Proof of Proposition 4

In the case of fixed costs the (**) curve in figure 2 is described by: $f = (1-p)\,w\,a_0$. A rise in R shifts the (*) curve downward and thus reduces both w and f.

In the case of linear costs the (**) curve is horizontal and is given by:

$$(***)\qquad \frac{\alpha}{1-\alpha}\,(f + R + E) + \frac{f}{\varepsilon} = \frac{4\,f^2}{9\,B\,R\,(1-p)}\,.$$

Hence an increase in R raises f. Q.E.D.

Proof of Proposition 5

In the case of fixed costs we can derive from equations (*) and (**) in the proof of Proposition 2 the following equation:

$$f = p\,(1-p)^{\frac{\alpha}{1-\alpha}}\, f^{\frac{-\alpha}{1-\alpha}}\, b - R - E\ ,$$

where b is a positive parameter. The R.H.S. is increasing in p if $1-p \geq \alpha$. Hence, a rise in $1-p$ reduces f.

In the case of linear costs examine equation (***) in the proof of proposition 4. It follows that f rises as $1-p$ rises. Q.E.D.

Appendix C
Raw Data For 1985

		GSP/ CAPITA	COST/ GLL	WRTOFF/ GLL	DEPO/ GLL	CNTR	BRNCH
1	Alabama	12,909	5.53	0.87	33.67	0.076	4.16
2	Arizona	15,217	5.16	0.62	31.04	0.196	14.25
3	Arkansas	12,681	5.90	1.24	28.10	0.009	2.72
4	California	18,850	5.48	1.50	27.90	0.037	10.91
5	Colorado	17,537	7.32	1.84	39.68	0.022	1.26
6	Connecticut	20,402	5.36	0.37	40.35	0.130	20.44
7	Delaware	17,630	3.17	1.12	7.14	0.060	6.17
8	Florida	14,462	6.21	0.75	33.42	0.023	6.15
9	Georgia	15,753	5.80	0.48	37.75	0.063	4.03
10	Idaho	12,975	5.53	1.20	21.84	0.212	11.50
11	Illinois	17,174	4.88	1.02	27.93	0.048	1.60
12	Indiana	14,593	5.13	0.83	28.91	0.020	4.47
13	Iowa	14,613	6.00	3.00	29.66	0.006	1.87
14	Kansas	16,482	5.51	1.85	28.56	0.007	1.35
15	Kentucky	13,739	5.06	0.85	30.65	0.038	3.32
16	Louisianna	17,771	6.03	2.00	32.11	0.016	4.08
17	Maine	13,645	6.30	0.26	29.01	0.107	20.08
18	Maryland	16,067	5.69	0.78	33.01	0.066	14.21
19	Massachusetts	18,242	5.61	0.47	33.62	0.047	16.55
20	Michigan	15,814	5.41	0.64	28.17	0.029	6.80
21	Minnesota	16,981	4.61	1.15	26.87	0.048	1.50
22	Mississippi	11,790	5.87	1.02	32.07	0.062	6.27
23	Missouri	15,734	5.49	1.14	37.18	0.017	1.97
24	Montana	13,992	6.08	2.09	25.81	0.014	1.15
25	Nebraska	15,975	6.52	3.10	33.44	0.023	1.42
26	Nevada	19,123	5.84	0.87	26.75	0.213	12.94
27	New-Hampshire	16,602	4.93	0.22	21.96	0.030	6.02
28	New-Jersey	18,821	5.36	0.37	40.03	0.026	17.13
29	New-Mexico	16,462	6.15	1.24	27.86	0.042	3.91
30	New-York	18,938	4.30	0.63	27.00	0.027	23.38
31	North-Carolina	14,983	4.97	0.28	28.34	0.103	30.78
32	North-Dakota	15,657	5.67	1.88	26.48	0.013	1.70
33	Ohio	15,599	5.16	0.72	26.29	0.023	8.76
34	Oklahoma	15,379	6.33	2.77	31.22	0.016	1.38
35	Oregon	14,491	5.63	0.75	24.72	0.225	10.55
36	Pennsylvania	14,582	4.64	0.47	25.07	0.023	10.40
37	Rode-Island	14,437	5.22	1.01	21.40	0.322	17.44
38	South-Carolina	12,543	6.88	0.49	38.09	0.137	11.76
39	South-Dakota	13,131	5.68	3.41	6.89	0.207	2.25
40	Tennessee	14,172	6.01	1.10	31.30	0.035	4.62
41	Texas	18,770	4.19	1.49	28.42	0.010	1.19
42	Utah	14,086	5.68	0.74	28.54	0.157	7.56
43	Vermont	14,794	4.53	0.42	19.78	0.082	9.76
44	Virginia	16,726	4.99	0.47	24.88	0.073	10.46
45	Washington	16,279	5.86	0.98	24.39	0.123	11.08
46	West-Virginia	12,153	5.76	0.67	26.96	0.009	1.87
47	Wisconsin	15,225	5.53	0.79	29.64	0.018	2.04
48	Wyoming	25,053	7.27	3.20	30.79	0.019	1.04
1985 Average		16,175	5.54	1.15	28.68	0.069	....

	Yearly Averages			
	GSP/ CAPITA	COST/ GLL	WRTOFF/ GLL	Interest Differential
1984	15,866	5.48	0.88	1.2
1985	16,175	5.54	1.15	1.5
1986	16,538	5.53	1.38	1.5
1987	16,979	5.66	1.32	1.1
1988	17,142	5.62	1.20	1.4
1989	18,313	5.63	1.26	1.6
1990	17,277	5.96	1.49	1.7
1991	16,332	6.36	1.60	2.5

Figure 1

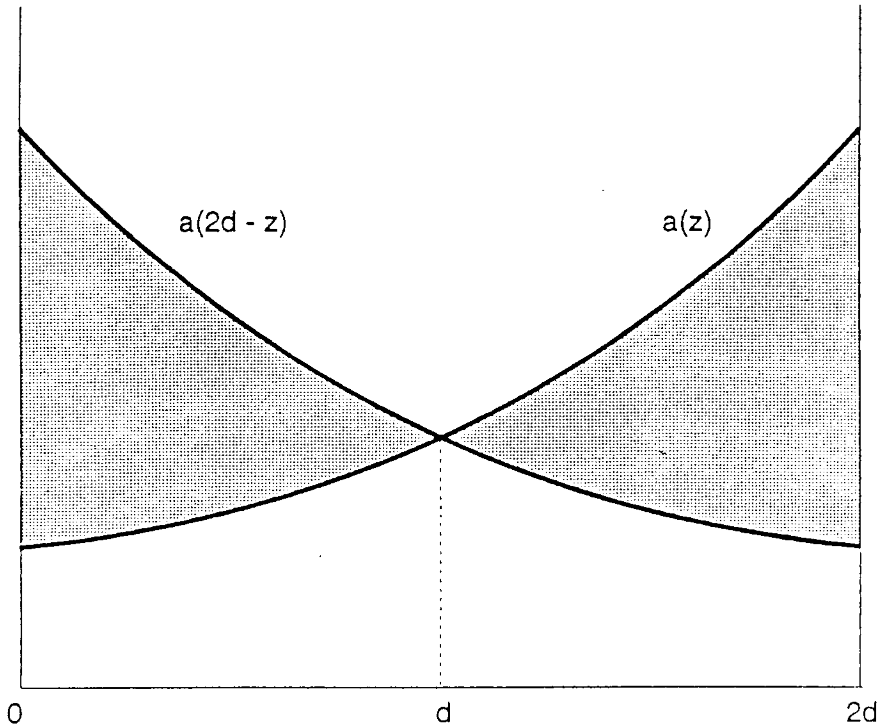

Figure 2

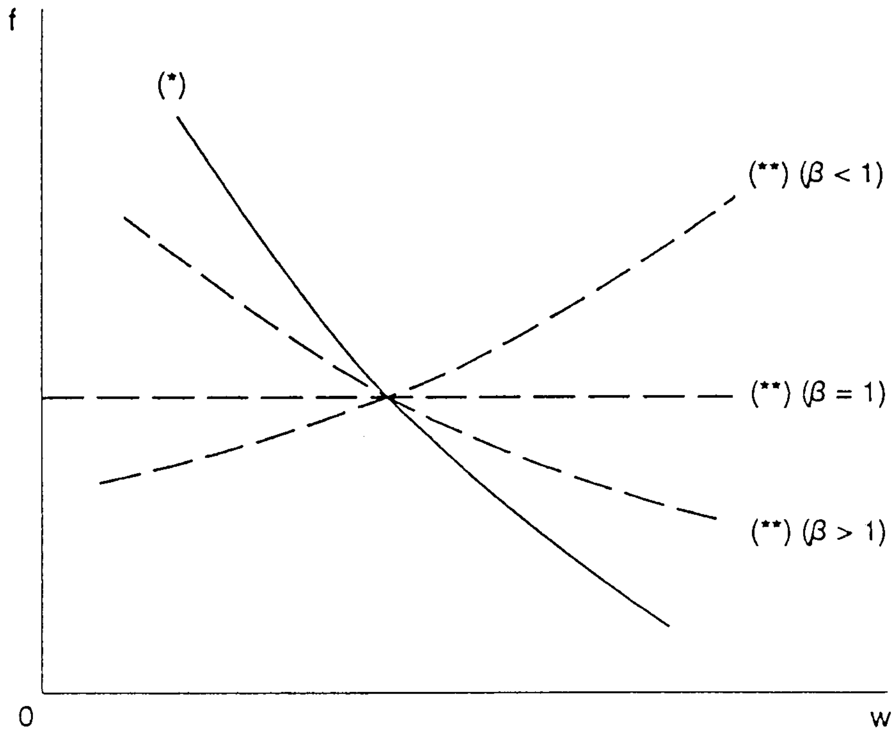

Figure 3. The Cost of Banking, 1985

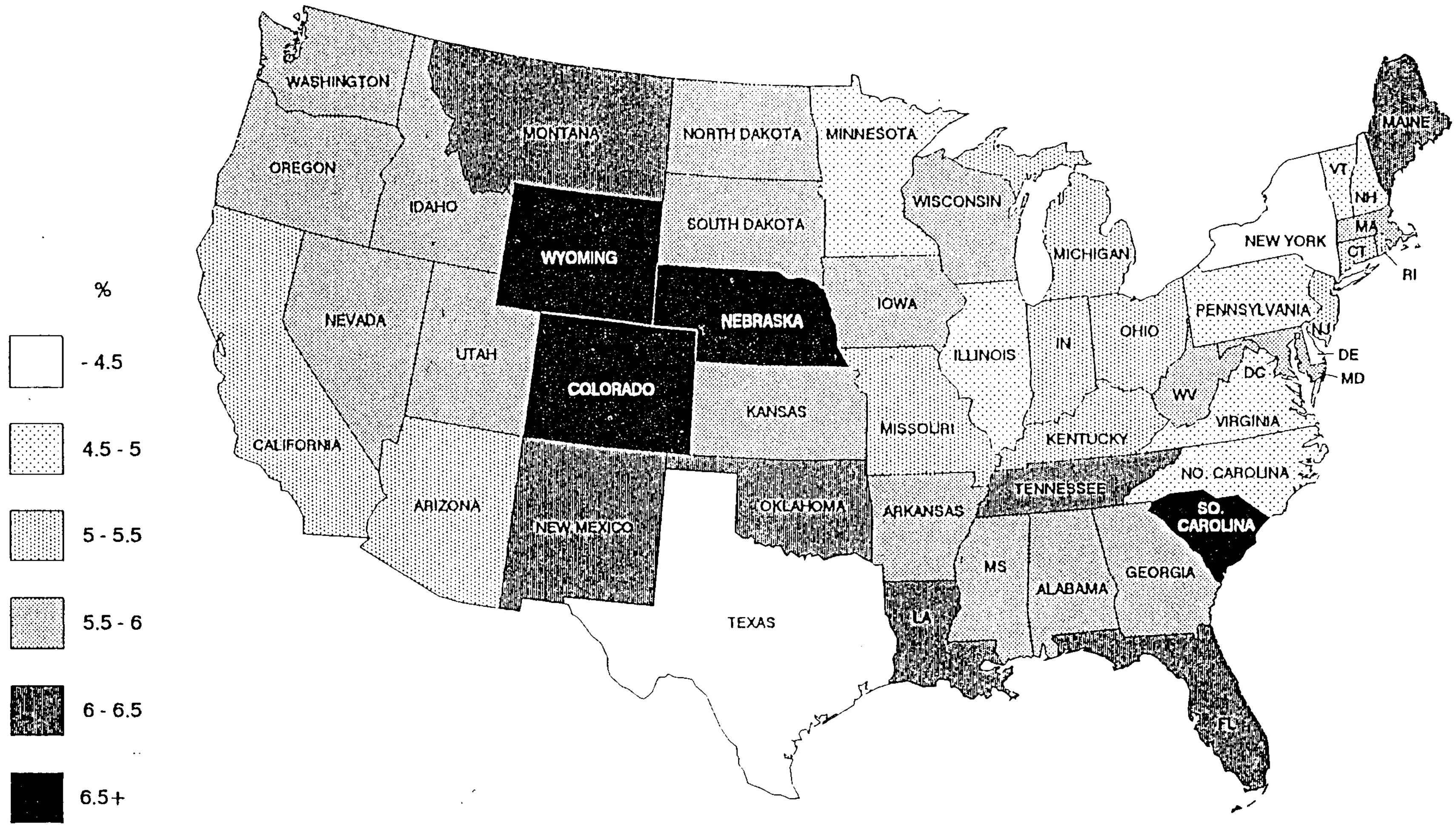

Figure 4
Financial Distress in Texas, 1984-1991

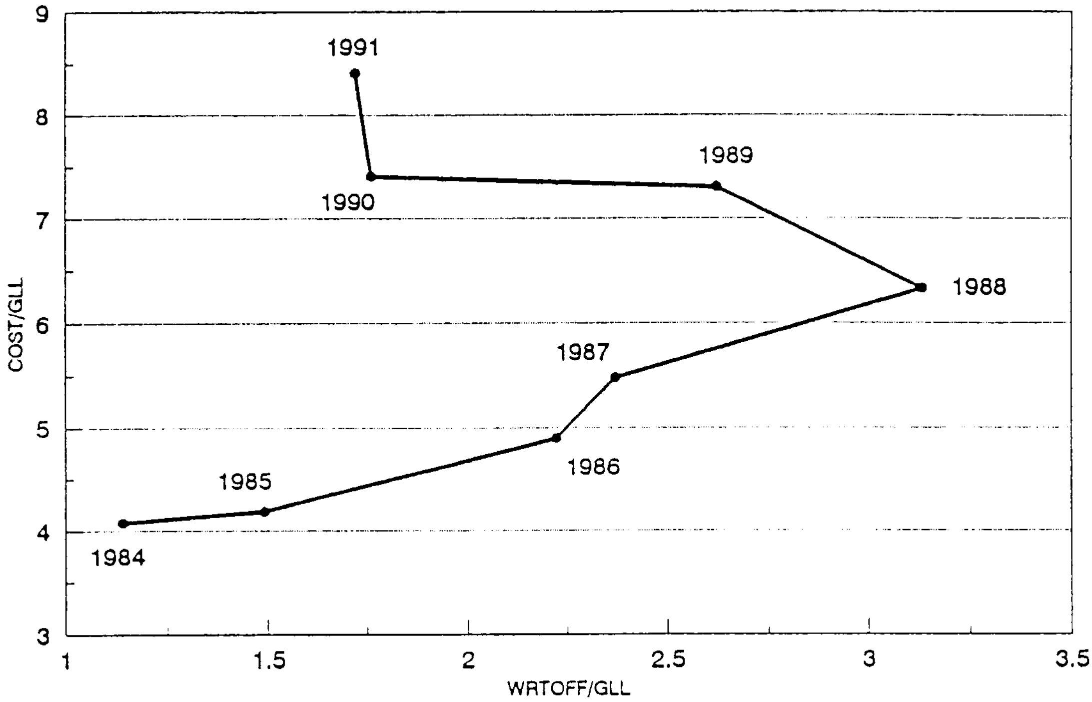

Figure 5
Year Dummies, the Panel, 1984-1991