Finance and Growth: Theory and New Evidence

Paul Harrison
Federal Reserve Board

Oren Sussman
Institute of Finance and Accounting
London Business School

Joseph Zeira
The Hebrew University of Jerusalem and CEPR

July 1999*

Abstract

This paper describes a feedback effect between real and financial development. The paper presents a new variable, which we call the cost of financial intermediation, through which the feedback between finance and growth operates. The theoretical part of the paper describes how specialization of financial intermediaries leads to such a feedback effect. The main result of the feedback is that differences in productivity across countries are amplified by financial intermediation. The empirical part of the paper uses US cross-state data from banks' income statements to measure the cost of financial intermediation and to provide evidence for the feedback effect between finance and growth.

JEL Classification: E44, G21, O16.
Key Words: Cost of Financial Intermediation, Economic Growth, Banks, Financial Development.

Mailing Address
Paul Harrison
Federal Reserve Board
Division of Research and Statistics
Washington DC 20551
USA
Email: paul.harrison@frb.gov

* We would like to thank Alex Cukierman, Zvi Eckstein, Stanley Fischer, Oded Galor, Saul Lach, Yoel Naveh, Steve Oliner, and Nathan Sussman for help in earlier drafts. We are grateful to the Chilevich Foundation in the Hebrew University for financial support. The views expressed herein are those of the authors and do not reflect those of the Federal Reserve Board or its staff.

Finance and Growth: Theory and New Evidence

1. Introduction

This paper presents two contributions to the literature on the interplay between financial development and economic growth. The first contribution is theoretical. We present a model of banking and growth, which generates a feedback effect between financial and real growth. As the economy grows, the banking sector becomes more specialized and thus more cost-effective, which feeds back to capital accumulation and growth. The second contribution is empirical. Using cross-state US banking data, we provide evidence in favor of the above effect. We show that the cost of financial intermediation is indeed decreasing with the level of the state's output. A by-product of the empirical analysis is an estimation of the cost of *ex-post* monitoring. It is strikingly high, more than 50 cents per dollar written off the balance sheet.

The theoretical part of the paper presents an information-based model of a monopolistically competitive banking system. It is based on Townsend's (1979) and Gale and Hellwig's (1985) theory of debt contracts, on Diamond's (1984) theory of banks as delegated monitors, and on Salop's (1979) model of spatial competition. In our model financial intermediation economizes the costs of monitoring ex-post returns of defaulting projects. We add a crucial assumption that monitoring costs are increasing with the distance between the bank and the project. This allows for banks' specialization, which generates a feedback effect in the following way. Economic growth increases banks' activity and profits, and thus induces entry of more banks. This entry reduces the average distance between banks and borrowers, promotes regional specialization and reduces the cost of intermediation. This in turn increases investment and economic growth. We then show that this feedback effect magnifies the effect of productivity on output. Hence, our model indicates that financial systems can account for some of the large international differences observed in output per capita.[1]

We refer to the negative effect of output on the cost of financial intermediation as the specialization effect. This effect is accompanied by another one, which operates in the opposite direction. Since monitoring is intensive in labor input, and since economic

[1] See Lucas (1990).

growth raises wages, it also tends to increase the cost of financial intermediation. We refer to this mechanism as the wage effect. It is important to recognize that there is no *a-priori* reason to believe that the wage effect is insignificant. Indeed, financial intermediation is a service, and the relative price of most services tends to increase with growth.[2] Our empirical findings show, however, that the wage effect is dominated by the specialization effect. It points out at least one respect in which financial services are 'special' relative to other services.

Our empirical findings support our choice of the cost of financial intermediation as a key variable, as it appears to be a significant one. The cost of banking is found to be rather high: 5-6% on average of banks' extended credit. The cost of ex-post monitoring is also estimated in the paper and is surprisingly high: more than \$0.50 on each \$1 of debt written off the balance sheet. Note, that this last result is important beyond the context of economic growth, as it provides evidence for the importance of costly ex-post monitoring in financial analysis.

Our paper fits into the current trend of research on finance and growth.[3] One characteristic of the new vintage of theoretical contributions is their attention to micro detail. Thus, for example, De la Fuente and Marin (1996) study the financing of technological innovations. Acemoglu and Zilibotti (1997) study the interrelations between project size, risk exposure and capital accumulation. Sussman (1999) studies the effect of financial innovations on risk exposure and saving. An important characteristic of recent empirical papers is their strong link with the empirical finance literature. Thus, for example, Rajan and Zingales (1998) show that industries with high dependence on external finance grow faster in financially developed economies. Levine and Zervos (1998) extend the finance-development analysis beyond the banking industry demonstrating the importance of *liquid* stock markets. Demirguc-Kunt and Maksimovic

[2] Kravis and Lipsey (1983) is the classic reference.

[3] The finance-development literature started with Gurley and Shaw (1955), Cameron (1967), Goldsmith (1969), Shaw (1973) and Kindleberger (1974), among others. Renewed interest has resulted from the new growth literature, which was pioneered by Baumol (1986), Romer (1986) and Lucas (1988) and has been growing rapidly since. The early contributions to the second phase of the finance-development literature include Greenwood and Jovanovic (1990), King and Levine (1993a, b), Sussman (1993), and Zilibotti (1994). See Gertler (1988), Pagano (1993), and more recently Levine (1997) for literature surveys.

(1998) show that firms can grow beyond the limits imposed by internal finance in an environment that respects legal norms and has active stock markets.

Our paper focuses on the banking industry. The theoretical part clearly defines the mechanism by which financial intermediation amplifies productivity and affects growth. Unlike other empirical papers in this area, we use *cost*, rather than *volume* to measure the level of development of the banking system.[4] This has a clear theoretical advantage: it is the cost of intermediation that directly affects the cost of capital and thus investment, while financial volumes may be affected by organizational features of the banking system that might not be related to the cost of capital. Another special feature of our paper is the use of US banking data, which is, presumably, much more reliable and more unified in terms of measurement and accounting standards than international data sets. To the best of our knowledge, the only paper that uses US data to test the 'finance-growth nexus' is Jayaratne and Strahan (1996). But theirs is more of an event-study approach: they show that states tend to grow faster after branch deregulation. Towards the end of their paper they argue that branch deregulation has affected growth through the effect it had on the 'quality of credit.' We take this approach several steps further, up to the point of direct estimation of the cost of ex-post monitoring.

The paper is organized as follows. Section 2 presents the theoretical model and Section 3 derives the equilibrium. Section 4 derives the feedback between the two sectors. Section 5 describes the data and the main empirical results and Section 6 discusses the endogeneity problem. Section 7 summarizes the paper. Appendix A contains proofs and Appendix B discusses the strength of the specialization effect.

2. The Model

Consider a small open economy with one physical good, which is used for both consumption and investment. Time is discrete and there are overlapping generations of two types of agents: workers and potential entrepreneurs. Each generation consists of a continuum of workers of size L and a continuum of potential entrepreneurs of size λL.

[4] Typically, M2 or bank credit over GDP is used in order to measure of development of the banking system. One of the few exceptions is the work by Faini et al (1993) that uses regional banking data from Italy.

Agents live for two periods each, they consume in the second period of life and are all risk neutral. A worker supplies one unit of labor inelastically in first period of life. A potential entrepreneur can become an entrepreneur, invest in a project in the first period of life, operate it in the second period of life and derive utility

$$U = c - E, \tag{1}$$

where c is consumption in second period of life and E is disutility from effort. Utility of a potential entrepreneur, who does not enter, is zero.

We next turn to describe projects and production. Each project is operated by one entrepreneur and requires both capital and labor inputs. One unit of capital is invested in each project one period ahead of time and it fully depreciates after one period of production. Projects differ by location and by productivity and both are random variables. As for location, projects are uniformly distributed along a circle with a circumference of length L. Note that location may represent many aspects in addition to geographical, like sectoral or technological. As for productivity, projects have a two state distribution, namely they can either succeed or fail. In case of success the project produces the following amount:

$$y = g(l) + a, \tag{2}$$

where y is output, l is labor input, g is a standard production function and a is a country specific productivity parameter.[5] In case of failure the project has zero productivity. The probability of success is p and we assume that productivity and location are independent of one another. We also assume that they are independent across projects.

We next describe the precise timing and information assumptions of the model. Every young potential entrepreneur in period t decides whether or not to make the required effort and become an entrepreneur. Once a positive decision is made, the entrepreneur is assigned a project, the location of which becomes common knowledge. Then one unit of capital is invested in the project, which must be financed with external funds. In the beginning of the next period productivity is realized. If the project is

[5] We assume for simplicity that the productivity parameter is additive rather than multiplicative. The qualitative results in the case of a multiplicative parameter are the same.

successful the entrepreneur hires labor and produces output y, and if the project fails the entrepreneur does not hire labor and does not produce. The realization of productivity is private information and is known to the entrepreneur only, including the information on employment, which is a perfect signal of productivity. Still, this information can be acquired by outsiders in an act of 'monitoring', which is costly.

Since monitoring is costly it is done by special firms, which operate as delegated monitors for many lenders, as described by Diamond (1984), and we call these firms banks. The monitoring technology requires both capital and labor and it depends on location as well. Banks are set up by investing a non-infinitesimal amount of capital B. After this amount is invested the bank is located and then it finances projects. In the next period the bank collects repayments, monitors failing projects and then its capital fully depreciates. Monitoring requires an amount $b(z)$ of labor per project, where z is the distance along the circle from the project of the bank. We assume that b is increasing in z. This assumption leads to the specialization effect.

We assume that banks have some monopoly power over borrowers. The "lending game" is the following: all banks, who compete for a certain project, simultaneously submit a 'take it or leave it' offer to the entrepreneur, who terminates the game by choosing the best offer. It turns out that the Nash equilibrium of this game leaves a considerable amount of bargaining power to the bank, as shown below.

Unlike the credit market, the deposit market is perfectly competitive. Banks raise finance through one period riskless deposits. We assume that there is full capital mobility and the domestic deposit market is fully integrated in the world's market.[6] Assume that the world's interest rate is constant and equals r. We denote the gross interest rate by $R = 1 + r$. Note that although deposits are internationally tradable, financial intermediation is nontradable, as banks must be located close to projects in order to economize on monitoring costs. Finally, assume that expectations are rational.

3. Equilibrium

3.1. Financial Intermediation

The financial structure of the economy emerges endogenously. Consider first the lender-entrepreneur contract. Obviously, the lender has to monitor the project sometimes, otherwise the entrepreneur will always declare failure, but the lender economizes on this costly monitoring activity. Townsend (1979) has shown for a more general framework, that the optimal contract is the standard debt contract, where the bank monitors the project only when the entrepreneur declares failure, and charges a fixed repayment I when the entrepreneur declares success, to provide an incentive for the entrepreneur to declare the true realization of the project. It can be shown that this is the optimal contract in our model as well, if random monitoring strategies (mixed strategies) are ruled out.[7] This debt contract leads to financial intermediation as in Diamond (1984). We assume that the parameters of the model are such that wages are small relative to investment in one project, and hence there are many lenders to each project. To economize on monitoring by many lenders and to prevent free-riding, monitoring is delegated to a single agent, which we call a bank. Moreover, by pooling together many independent projects, the bank avoids agency problems with its depositors.[8]

[Insert Figure 1 here]

Denote the number of banks who enter in period t by N_t. This is a finite number due to the lumpiness of the setup cost of a bank. Assume that banks are located on the circle in equal distances from each other. Let $2d_t$ be the distance between each two neighboring banks, namely $2d_t = L/N_t$. Consider now an entrepreneur which borrows from a bank in period t. The amount of repayment set in the debt contract I_t depends on the distance between the project and the two closest banks. Figure 1 shows how much labor two neighboring banks need in order to monitor a defaulting project, as a function of the project's location. It is evident that each bank has a cost advantage within a distance of d_t, and can therefore undercut any offer made by its competitor. Hence, in the

---

[6] The analysis of a closed economy version of the model, where the interest rate is endogenously determined, yields similar results. The open economy setup is more suitable to our data set of the US.

[7] Random monitoring is often ignored in such models, since it usually yields an optimal contract which is very different from the standard debt contract (see Mookherjee and Png, 1989). It can be shown that in this model random monitoring would further strengthen the feedback effect. Hence, ruling out mixed strategies does not change the main results of the paper.

[8] Note, that banks' ability to precommit must be assumed as well, since banks monitor all defaulting for deterrence reasons.

Nash equilibrium finance is supplied by the bank with the cost advantage, according to the most favorable terms, which the competitor can offer. Therefore, the equilibrium repayment for a project at distance z from the bank, $0 \le z \le d_t$, is

(3)
$$I_t(z) = \frac{R}{p} + \frac{1-p}{p} w_{t+1} b(2d_t - z).$$

Note, that the bank acts as a perfect price discriminator over the projects it finances. It extracts a high monopoly rent from projects located near to it, but breaks even on projects on the margin of its market segment.

3.2. Real Investment

Agents make their entry decision, namely whether to become entrepreneurs or not, before location and productivity are realized. In equilibrium the disutility from effort is equal to net expected profit given the perfectly forecasted next period wages. Namely:

(4)
$$p\left[a + \pi(w_{t+1}) - \frac{1}{d_t}\int_0^{d_t} I_t(z)dz\right] - E = 0,$$

where the function π is defined by

(5)
$$\pi(w) = \max_l [g(l) - wl],$$

and is a decreasing function of wages. The expected costs of financing a project are

(6)
$$\frac{p}{d_t}\int_0^{d_t} I_t(z)dz = R + (1-p)w_{t+1}\Psi(d_t),$$

where Ψ is given by:

(7)
$$\Psi(d) = \frac{1}{d}\int_0^{d} b(2d - z)dz,$$

and is increasing with d. The equilibrium entry condition for entrepreneurs is, therefore:

(8)
$$p[a + \pi(w_{t+1})] - R - (1-p)w_{t+1}\Psi(d_t) - E = 0.$$

3.3. Financial Investment

We now turn to describe the entry condition of banks. Let K_t be the amount of capital invested in real production in period t, which also equals the number of projects initiated. Let $k_t = K_t/L$ denote the density of projects on the circle, which is also the capital-labor ratio in the real sector. Expected profits of a bank consist of repayments from solvent projects minus monitoring costs of defaulting projects and minus payments for deposits. The entry condition for banks in period t is therefore:

(9) $$BR = 2pk_t \int_0^{d_t} I_t(z)dz - 2(1-p)k_t w_{t+1} \int_0^{d_t} b(z)dz - 2k_t d_t R = (1-p)k_t w_{t+1} \Phi(d_t),$$

where Φ is:

(10) $$\Phi(d) = 2\int_0^{d} [b(2d-z) - b(z)]dz.$$

Note that equation (9) has a simple interpretation. The left hand side is the bank's setup cost, while the right hand side is the bank's monopolistic rent, where Φ captures the bank's cost advantage in labor units in its market segment, namely the area between the curves in Figure 1. Note, that since Φ is increasing, the distance d_t and the derived number of banks N_t are uniquely determined by equation (9).

3.4. Labor Market Equilibrium

As we see above, entry to both the real and the financial sector depends on future anticipated wages w_{t+1}. Next, we present the equilibrium condition in the labor market in period t+1 in order to determine the wage rate. Profit maximization (5) leads to the standard first order condition: $g'(l) = w$, from which the demand for labor per project $l^d(w)$ is derived. The aggregate demand for labor in the real sector in period t+1 is $pK_t\, l^d(w_{t+1})$, and the aggregate demand for labor in the financial sector is $(1-p)N_t\, k_t 2\int_0^{d_t} b(z)dz$. In equilibrium the sum of the two demands equals supply L. The labor market equilibrium condition in period t+1 is, therefore:

(11) $$pk_t\, l^d(w_{t+1}) + (1-p)k_t\, \Theta(d_t) = 1,$$

where Θ measures the average labor requirement of monitoring:

(12) $$\Theta(d) = \frac{1}{d}\int_0^d b(z)dz.$$

Equation (11) uniquely determines the equilibrium wage rate as a function of investment in the real sector k_t and in the financial sector, represented by d_t.

3.5. General Equilibrium

The equilibrium in period t is characterized by three variables: the density of capital in the real sector k_t, the distance between banks d_t, and the equilibrium wage rate in the next period w_{t+1}. The equilibrium is, therefore, determined by three equilibrium conditions: the entry condition to the real sector (8), the entry condition to the financial sector (9), and the labor market equilibrium condition (11). This is shown in the following proposition.

Proposition 1: If the financial sector is sufficiently small, namely if B and $1\text{-}p$ are small, there exists a unique and stable equilibrium. The equilibrium is time invariant.
Proof: In the Appendix.

4. Output and the Cost of Financial Intermediation

We next introduce a new variable which plays a key role in the following analysis.

Definition: The cost of financial intermediation in period $t+1$, f_{t+1}, is the average cost of the bank per unit of lending. It includes setup and monitoring costs, but not the cost of deposits. Formally:

(13) $$f_{t+1} = \frac{BR + 2(1-p)w_{t+1}k_t\int_0^{d_t} b(z)dz}{2k_t\, d_t}.$$

Before we analyze the role of the cost of financial intermediation in economic growth, we rewrite equation (13) by substituting condition (9) of the equilibrium entry to the financial sector. Deleting all time subscripts, since the equilibrium is time invariant, we get:

$$f = (1-p)w\Psi(d)\,. \tag{14}$$

This equation reveals the two main effects of economic growth on the cost of financial intermediation. On the one hand, growth raises wages and that has a direct positive effect on the cost of intermediation. We call this the wage effect. On the other hand, growth increases entry of banks and reduces the distance d. That has a negative effect on the cost of intermediation, which we call the specialization effect.

We now formalize the effect of economic growth on the cost of financial intermediation by deriving the equilibrium distance between banks. By substituting the banks' entry condition (9) in the labor market equilibrium condition (11) we get:

$$pl^{d}(w) + (1-p)\Theta(d) = \frac{(1-p)w}{BR}\Phi(d)\,. \tag{15}$$

This equation defines a function: $d = d(w)$, which describes how specialization depends on the level of wages, namely on economic growth. Note that this function is decreasing, if equilibrium exists.[9] This is the specialization effect: as the economy grows and wages increase, more banks enter and become more specialized over a smaller market segment.

In order to see both the wage effect and the specialization effect on the cost of financial intermediation we substitute $d(w)$ in equation (14) and get:

$$f = (1-p)w\Psi[d(w)]\,. \tag{16}$$

Equation (16) displays clearly the ambiguous effect of economic growth on the cost of financial intermediation, as it shows two opposite effects: the direct positive wage effect, and the indirect negative specialization effect. The ambiguity is reflected in the slope of the *ff* curve, which is derived from equation (16) and is plotted in Figure 2 in the (w, f) plane. The curve is upside sloping when the wage effect dominates and is downward sloping when the specialization effect dominates, as shown by ff_1 and ff_2, respectively.

[9] Plot the LHS and the RHS of equation (15) against d. Both curves are increasing, but at d=0 the LHS is positive while the RHS is equal to 0. Hence the LHS intersects with the RHS from below, if an equilibrium exists. The analysis of the effect of w is straightforward.

The cost of financial intermediation is not only affected by economic growth, as shown in equation (16), but it affects it as well, through net profits of entrepreneurs. To see this, substitute equation (14) in the entry condition to the real sector (8) and get:

$$p[a+\pi(w)]-R-f-E=0. \tag{17}$$

From this equation we deduce that if the cost of financial intermediation rises, entrepreneurial profits fall and so do investment and output. Equation (17) is described by a second curve in the plane (w, f), which we denote pp. The pp curve is downward sloping, since a lower f increases entry of new entrepreneurs and that raises output and wages. Note, that pp intersects with the ff curve from above when the equilibrium is stable.

[Insert Figure 2 here]

The following proposition states the main theoretical result of the paper. It shows that the feedback between the real and the financial sector depends on which effect dominates, the specialization effect or the wage cost effect.

Proposition 2: If the specialization effect dominates, the effect of productivity on output and wages is amplified. If the wage cost effect dominates, the effect of productivity is dampened.

Proof: A rise in a shifts the pp curve upward and to the right. In an economy without financial intermediation $f=0$ and the rise in a increases w by

$$-\frac{1}{\pi'(w)}.$$

If the specialization effect dominates, as described by the curve ff_2, the wage rate w rises by more, namely by

$$\frac{dw}{da}=-\frac{1}{\pi'(w)-\frac{1}{p}\frac{\partial}{\partial w}\{(1-p)w\Psi[d(w)]\}}>-\frac{1}{\pi'(w)}.$$

It is clear that in this case the cost of financial intermediation f declines as output and wages increase. Proof of the case of dominant wage cost effect is similar.

QED

If differences across countries are due to differences in productivity of the real sector, namely due to changes in a, then the sign of the correlation between output and the cost of financial intermediation depends on which effect dominates. If the specialization effect dominates, the correlation is negative, due to the negative slope of the *ff* curve, while if the wage cost effect dominates, the correlation is positive, due to the positive slope of the *ff* curve. This is an important empirical implication of the model. Note, though, that differences between countries can be a result not only of differences in real productivity a but of differences in banking technology, which shift the *ff* curve. Consider for example the effect of a financial liberalization, which can be viewed in our model as reducing entry costs to banks, namely reducing B. Such a liberalization reduces d and shifts the *ff* curve downward. Hence, wages increase and the cost of financial liberalization declines. As a result, changes in banking technology also lead to a negative correlation between output and the cost of financial intermediation.[10] This observation is important to the empirical analysis in the paper, since we need to be careful in analyzing the correlation between w and f. We need to examine whether the *ff* curve is downward sloping or whether most of the shocks are from the banking technology.

As shown above, the slope of the *ff* curve is crucial to our understanding of the effect of financial intermediation on economic development. We next claim that the specialization effect weakens with output. To see this consider the following heuristic argument. If output is large and much capital is invested, many banks enter and the distance between them d becomes small. As a result $\Psi[d(w)]$ becomes less responsive to changes in w and the specialization effect is reduced. Hence, we would expect the specialization effect to dominate for economies with low output and the wage effect to dominate for economies with high output. This is indeed shown by use of specific production and monitoring functions in Appendix B. Which effect dominates becomes an empirical question, as we do not know at which level of output the specialization effect is

[10] Empirical evidence for this is presented in Jayaratne and Strahan (1996) where a financial liberalization contributes to real growth.

overcome by the wage effect. Indeed, this empirical question is analyzed in the following sections of the paper.

5. Data and Empirical Results

In this section we introduce cross-state US data in order to test the model of economic growth and financial intermediation developed above. The main conclusion we draw from the data is that economic growth reduces the cost of financial intermediation, namely that the specialization effect dominates the wage effect. We also show that the cost of ex-post monitoring is quite high, as the costs of handling bad debt by the bank are more than 50% of the defaulted debt.

5.1. Description of Data and Variables

The data we use, of US banking, is a large improvement over international data, since measurement errors are less of a concern and accounting and statistical procedures across states are unified. The state banking data are more detailed than most international sources, which allows for better control of omitted factors. Using it also enables us to match the data with a rich set of state-level demographic variables. The banking data are from Federal Deposit Insurance Corporation (FDIC) publication of state consolidated balance sheets and income statements for commercial banks from 1965-1995. This includes the 48 mainland states, excluding Alaska, Hawaii, and DC. Two other outliers, Delaware and South Dakota (which are both exporters of banking services), are included. The results that follow are robust to their exclusion. The per-capita state output data are taken from the Bureau of Economic Analysis from 1982-1994. The merged data are a panel of 48 states from 1982 to 1994.

The main relationship we explore in the empirical analysis is that between output per capita and the cost of financial intermediation, in order to determine whether the specialization or wage effect dominates. To do so we examine the following relationship, which is an empirical linearized restatement of equation (16) of the model:

$$\text{(18)} \qquad \text{COST/GLL} = \alpha 1 + \alpha 2\ \text{WRTOFF/GLL} + \alpha 3\ \text{DEPO/GLL} + \alpha 4 \log(\text{GSP/CAPITA}) + e.$$

The variable COST/GLL measures the cost of banking per \$1 of credit, where COST is banks' operational expenses (labor, fixed capital, and other expenses) not including interest expenses, and GLL is extended credit or "gross loans and leases". Note, that the cost of banking is not identical to the cost of financial intermediation f, since banks produce more services than intermediation, such as clearing and liquidity. To control for this, DEPO/GLL is added as an independent variable, where DEPO is the amount of non-interest bearing deposits, and it proxies for the prevalence of non-intermediation services. The probability of default $1\text{-}p$ is proxied by WRTOFF/GLL, where WRTOFF is the amount of bad debt written off the balance sheet. The variable w in the model measures both the wage rate and also the level of output per capita. We proxy this variable by output per capita, namely by GSP/CAPITA, where GSP is real gross state product (the state equivalent to GDP for a country) in 1992 prices and CAPITA is the state's population. The deposit interest rate R is not included in the empirical analysis, as the model suggests that it is equal across states. The empirical analysis, therefore, aims at finding the sign of the coefficient $\alpha 4$.

[Insert Table 1 here]

Table 1 provides summary statistics of the main variables over the sample period used for estimation: 1982-1994. During this period the cost of banking has been strikingly high and averaged 5.85 cents per \$1 credit. This cost has large variation across states, as shown in figure 3. The magnitude of the cost of banking can be better appreciated against another variable, bank's income per dollar of extended credit. Over 1982 to 1994 banks earned on average 14.8 cents in interest and fees on each dollar loaned out. Hence, more than a third of bank income has been eaten up by operational expenses. Figure 4, which plots the cost and income variables over time, shows that operation costs have recently been above 50 percent of banks' income. Also of particular interest is the realized probability of default: 1.13% on average over 1982-1994.

[Insert Figures 3 and 4 here]

5.2. Estimation of the Model

We next turn to the estimation of our model. When trying to examine the effect of output per capita on the cost of intermediation, we must recognize the potential for causality both ways between the two variables. This problem is addressed by use of instrumental variables in the following section. In this section we present OLS regressions of the cost variable on output per capita, in order to get a preliminary understanding of its effect.

Table 2 reports the OLS estimation results of equation (18), as well as subsequent refinements. Year and state dummies are included and jointly significant, but are not reported. Absolute values of t-statistics are shown in parentheses using White heteroskedasticity-corrected standard errors, "obs" denotes the total number of observations, "χ^2" reports the Breusch-Godfrey chi-square test (Greene, 1990, pages 426-429) for serial correlation in the residuals (with the critical value for rejecting no serial correlation at the 5 percent error level shown in parentheses), and "R^2" is the adjusted R-squared statistic.

[Insert Table 2 here]

As shown in Table 2, estimation of equation (18) reveals a significant amount of residual autocorrelation. Hence, our baseline regression, equation (19), modifies that equation by adding the lagged dependent variable to control for the autocorrelation:

$$(19) \quad \text{COST/GLL} = \alpha 1 + \alpha 2\ \text{WRTOFF/GLL} + \alpha 3\ \text{DEPO/GLL} + \alpha 4 \log(\text{GSP/CAPITA}) + \alpha 5\ \text{lag(COST/GLL)} + e$$

Equation (19) can be justified not only via econometric considerations, but also by the model itself. Note that in our model the effect of output on the cost of banking operates through the entry and exit of banks, which alters the degree of specialization. Since this process will take more than a year, the lagged cost variable could be added to control for the impact of past output changes on current costs. By including the lagged cost variable we also implicitly control for other factors, outside the model, that might similarly influence the level of costs for multiple periods. This implies that the long-run effects of the other variables, namely output and default rate, should therefore be greater than the estimated short-run effects.

5.3. The Relation between Growth and Costs

The results in Table 2 are consistent with the theoretical model and lend support to our main conjecture of specialization since they indicate that the coefficient of GSP/CAPITA is statistically significantly negative. From the estimation of equation (19), a one hundred percent increase in per-capita GSP is associated with an 8 tenths of a cent change in the cost of banking per dollar of extended credit. Because the average cost of banking per dollar of extended credit is 5.85 cents, this .8 cent effect translates into an elasticity of costs with respect to income of .13 – a one hundred percent increase in per-capita GSP would decrease costs by 13 percent. While this might seem to require a big change in per-capita GSP, keep in mind that the cross-sectional variation in GSP/CAPITA is large – with output in rich states (like California and Connecticut) almost doubling that in poor states (like Mississippi and West Virginia). Our conclusion is thus that economic development in poor states (and moreover countries) could significantly reduce the costs of banking.

5.4. The Cost of Monitoring Defaults

Also as expected, Table 2 reports that both the default rate and deposit rate have significant positive coefficients. The indication that the amount of bad debt has a strong positive impact on the cost of banking is evidence supporting the Diamond and Townsend assumptions of costly ex-post monitoring. Moreover, the effect of the default rate on banking costs is strikingly high. Every dollar of bad debt raises gross costs by more than 50 cents. This result indicates that ex-post monitoring in the case of default is not only qualitatively significant, but quantitatively as well.

This number is much higher than the estimates of bankruptcy costs that are cited in the corporate finance literature where such costs are usually thought to be relatively small and insignificant. Warner (1977) and Weiss (1990) estimate the direct costs of bankruptcy to be near 3 percent (of market value prior to default). Two important differences between these studies and ours should be stressed. First, we measure the cost imposed on the lender, rather than on the defaulting corporation. Second, they focus on large public corporations whose size and public status are not independent of their low bankruptcy costs. For small firms, the corporate finance literature estimates bankruptcy costs that are much closer to our numbers, perhaps as much as 25 percent of a firm's value.[11]

Furthermore, our estimated cost of monitoring, high as it appears, only explains a relatively small part of the total cost of banking. Recalling that banks wrote-off just over 1 percent of their extended credit (GLL), the default portion of costs amounts to less than .6 cents of the 5.85 cents total per dollar extended – or about 10 percent of the gross costs of banking.

5.5. Robustness to Omitted Variables

These results should be particularly robust because the fixed-effects estimation wipes out much of the potentially spurious variation, and instead relies on identification from the within state fluctuations over time. The fact that we are essentially exploiting less variation and still finding significant results makes the interpretation more compelling. The year effects will pick up common business-cycle phenomenon that influence both output and costs, and would otherwise lead to a spurious relation between those variables. The state effects similarly attenuate the GSP/CAPITA coefficient, as both output and the cost of banking vary with state specific factors that are unlikely to be causal. If not for the potential confounding influence of these state characteristics, a pure cross-sectional regression might provide a cleaner experiment for estimating our hypothesized effect within a development context. As expected, year-by-year cross-sectional regressions (which are not reported) find much stronger effects from output onto bank costs. However, the variation across state characteristics makes it difficult to believe that the estimated coefficient is properly identified in this setting.

Although the theoretical model does not allow for additional cross-state differences in banking behavior, in reality some states may house "financial centers" characterized by lower costs of banking. Cost reductions in these centers could derive from scale economies, greater competition among banks, or from borrower composition.[12] We attempt to control for variations of this sort by experimenting with

---

[11] See Grinblatt and Titman (1998, page 545).

[12] For instance, there may be more large borrowers in centers which may be cheaper to monitor (per dollar of loans) than smaller borrowers.

proxies for the scale of banking activity. Our preferred variable is GLL/GSP which measures the amount of extended credit relative to the size of the state economy. This variable should be larger in financial centers and smaller at the periphery. Thus, it also proxies for the amount of cross-state banking that that state receives. We therefore estimate equation (20), which is equation (19) supplemented by the scale variable:

$$(20) \quad COST/GLL = \alpha 1 + \alpha 2 WRTOFF/GLL + \alpha 3 DEPO/GLL + \alpha 4 \log(GSP/CAPITA)$$
$$\alpha 5 \text{lag}(COST/GLL) + \alpha 6 \log(GLL/GSP) + e$$

The results of the estimation of (20) are also presented in Table 2. The scale variable is found to be marginally significant and enters with a negative sign in the equation. This is further support for the specialization effect, namely that more banking activity reduces the cost of banking. The coefficients on the other independent variables are essentially unaffected by the additional variable. In equations without the lagged dependent variable the scale variable is much more significant. This is because adding the lagged cost variable is similar to controlling for financial centers (since they have lower costs yesterday and today). Alternative proxies for the amount of financial activity were tried with a similarly insignificant impact on the other coefficient estimates: GLL divided by the number of bank workers, GLL divided by the number of banks, the number of banks, as well as the number of branches.

We also attempted to account for the sectoral composition of the state, under the hypothesis that manufacturing firms are larger, more transparent, and thus cheaper to monitor. In the panel estimation the percentage of GSP in manufacturing has an insignificant coefficient estimate, although in cross-sectional regressions it comes out significantly negative. Since the intensity of manufacturing is slow to change over time, it provides no additional explanatory power once the state effects and lagged cost variable are included in the panel regression. Again this point highlights the credibility of our results and substantially mitigates concern over omitted variables. The year effects and the lagged cost variable should pick up business cycle variations while the state effects together with the lagged cost variable appear to pick up the ex-ante cost of monitoring that might otherwise show up in the scale and sector variables.

5.4. Robustness to Specification and Estimation

Diagnostic tests on the estimation are also encouraging. The estimation of equation (19), as designed, is found to be free from autocorrelation. However, because estimation of dynamic fixed-effects models is potentially biased we check the robustness of these results using a number of alternative approaches.[13] For example, the findings from random-effects estimation are similar to those already reported and again support a statistically significant negative output coefficient.[14] Furthermore, the results are not sensitive to a reduction in the number of cross-sectional dummy variables when states are combined by region, or to an increase when they are split into two dummy variables per state. Deleting individual years or states from the analysis also preserves the qualitative findings, as does estimation in first differences or percent changes. Finally, we note that all of our findings are unchanged, but the size of the estimated effects stronger in absolute value, if the lagged cost variable is excluded from the regressions (as in equation (18)).

6. Endogeneity and Instrumental Variables Estimation

The equations estimated in section 5, which examine the effect of output on banking costs, ignore the effect in the other direction, of costs on output, as implied by equation (17) in the model. One way to interpret our results is to assume that the banking technology is fixed across states and the only differences are in technology a. Namely, that the only curve which shifts in figure 2 is *pp*, while the *ff* curve is fixed. Although such an assumption is realistic, as evidenced below, we attempt in this section to isolate the effect of output on banking costs from the opposite effect by use of instrumental variables.

We first consider the time-series relationship between banking costs and output in more detail by examining how lags of each variable impact their current realization.

[13]See Nickell (1981), Hsiao (1986, ch. 4), and Baltagi (1995, ch.8). The asymptotic bias will tend to zero in our application. To establish finite sample robustness we proceed to show that the results are not sensitive to the size of N relative to T and thus that the potential bias is not a significant concern.

[14]Although the random effects estimates are qualitatively similar to the fixed effects estimates, it seems clear that we should prefer a fixed effects model because the constant differences between states are likely non-random.

Table 3 summarizes a standard "Granger Causality" analysis (see Hamilton, 1995, pages 302-309).

[Insert Table 3 here]

The results suggest that changes in GSP "cause" changes in COST, so that even when controlling for the impact of past COST on current COST, lagged GSP matters. To the contrary, lagged banking costs are not found to be an important determinant of current GSP (not when controlling for lagged GSP). In fact the lagged COST appears to enter with the "wrong" sign, although it is not significant. These results lend support to assuming that most of the action across states is due to differences in the real economy, namely shifts in *pp*, rather than differences in the banking technology, namely shifts in *ff*. But in order to be confident that we isolate the impact of output on banking costs we also worry about the contemporaneous endogeneity between these two variables, since our results could simply be picking up the effect of current banking costs on GSP rather than GSP on bank costs. To address this concern we use instrumental variables (IV) for the variable GSP/CAPITA and 2SLS to re-estimate the coefficient of interest, namely $\alpha 4$ in equation (19).

We use three different first-stage regressions based on three different, but plausible, sets of instruments for the level of GSP/CAPITA. The first set of instruments is simply lagged GSP/CAPITA. The second candidate is GSP/CAPITA in 1982. This can be viewed as an initial condition for output that is exogenous because it predates the first year we consider. The third IV equation uses demographic and regional variables from the early 1980's (again prior to the beginning of our sample) to forecast per-capita output.[15] This last approach follows many recent growth regressions, both across countries and across states in the US, as in Barro and Sala-i-Martin (1992), Sala-i-Martin (1997) and Sachs and Warner (1997). Using data prior to our sample period should ensure the exogeneity of the instruments, although we will perform Hausman-type tests to

[15] We use the following variables: regional dummies following the Bureau of Economic Analysis divisions of "New England", "Middle Atlantic", "Pacific", "Midwest", "South Atlantic", "East South Central", "West South Central", and "Mountain"; the state high-school completion rate in 1980; the mean state temperature; the state infant mortality rate in 1980; a state dummy variable for ocean access; and the state percentages of 1982 GSP from farming, from manufacturing, and from natural resource exploitation. These variables proxy for the productivity variable a from our model.

verify that the chosen instruments are indeed uncorrelated with the second-stage residuals.

The estimated first-stage equations are, therefore, all variants of:

(IV) $$\log(GSP/CAPITA) = c1 + \gamma \mathbf{Z} + \beta \mathbf{X} + u$$

where **Z** represents the three different sets of instruments and **X** is the vector of exogenous regressors. These equations yield forecasts of log(GSP/CAPITA), which we denote "GSP_IV", and which are used in the second stage regression to estimate the coefficient on output $\alpha 4$:

(21) $$COST/GLL = \alpha 1 + \alpha 2\ WRTOFF/GLL + \alpha 3\ DEPO/GLL + \alpha 4\ GSP\_IV + \alpha 5\ \mathrm{lag}(COST/GLL) + e.$$

The basic first and second stage regressions are summarized in the top panel of Table 4, with year dummies included but not reported. For each set of instruments the estimated effect of output on bank costs is significant and negative – the evidence in favor of the specialization effect is robust. For convenience the OLS estimation results of equation (19) from Table 2 are also included. The 2SLS point estimates of $\alpha 4$ are significantly different from the OLS estimates except in the case of the first IV equation, where the instrument -- lagged GSP -- is highly correlated with current GSP and yields a first-stage adjusted-R^2 of .975. It seems plausible that the 2SLS estimates for $\alpha 4$ would be attenuated, since the proposed endogenous effect is a negative impact from cost onto GSP, suggesting a spurious augmentation of the negative effect from GSP onto cost. This effect is indeed observed in Table 4, except for IV4, which uses IV estimates for the deposits and defaults variables in addition to GSP and thus could be controlling for a range of inter-relations.

[Insert Table 4 here]

To assess the plausibility of our instruments as exogenous we compute the standard over-identifying test by regressing the second stage residuals on the first stage instruments.[16] For each set of regressions we find that the instruments are exogenoeous.

[16] If the instruments and errors are truly exogenous the R^2 from this regression should be zero. The test statistic $T*N*R^2$ is distributed as a chi-square with degrees of freedom equal to the number of instruments in the first stage minus one.

However, in order to apply the over-identifying test to the first and second IV equations, which are otherwise exactly identified, their first stage model is augmented with the percent of output from manufacturing as an instrument (it is significant in both equations) to yield second-stage equations (IV 1a) and (IV 2a) in Table 4.

Of course, we can not rule out the possibility that the deposits and default variables are also endogenous, since bank costs might also feed back into these variables. To address this problem we run 2SLS instrumenting for DEPO/GLL and WRTOFF/GLL using lagged right-hand side variables and the demographic variables. The results appear in equation (IV 4) in Table 4. The estimates are quite similar to the OLS and 2SLS results, except that deposits are no longer significant, and still support the main conjecture that growth reduces the cost of banking. This regression also addresses the concern that our estimate of the ex-post monitoring costs of defaults is endogeneous with costs.

7. Summary

This paper suggests that we examine the relationship between the real and the financial sector by use of a new variable, the cost of financial intermediation. On the one hand, this variable is affected by the level of output, due to both the wage and the specialization effect. On the other hand, this variable reflects how the financial sector affects the real sector development. This variable, therefore, summarizes the causal effects from one sector to the other. Furthermore, this variable can be measured, though imperfectly, and thus enable us to perform empirical tests of the various effects between finance and growth.

When we apply our analysis to US data we show that, indeed, the specialization effect exists and is quite strong. We learn from this that there should be a feedback effect between the real and the financial sector in most countries, since the US has relatively high output per capita. This feedback effect amplifies primary differences in productivity across countries, and thus helps to explain international differences in output per capita.

References

Acemoglu, Daron, and Fabrizio Zilibotti, "Was Prometheus Unbound by Chance? Risk, Diversification, and Growth," *Journal of Political Economy* 105 (1997), 709-751.

Atje, Raymond, and Boyan Jovanovic, "Stock Markets and Development," *European Economic Review* 37 (1993), 632-640.

Baltagi, Badi H., *Econometric Analysis of Panel Data*, Chichester, UK, Wiley, 1995.

Baumol, William J., "Productivity Growth, Convergence, and Welfare: What the Long-Run Data Show," *American Economic Review* 76 (1986), 1072-1085.

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

Cameron, Rondo, *Banking and Development in the Early Stages of Industrialization*, New York, Oxford University Press, 1967.

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

De la Fuente, Anjel, and Jose Maria Marin, "Innovation, Bank Monitoring, and Endogenous Financial Development," *Journal of Monetary Economics* 38 (1996), 269-301.

Demirguc-Kunt, Asli, and Vojislav Maksimovic, "Law, Finance, and Firm Growth," *Journal of Finance* 53 (1998), 2107-2137.

Diamond, Douglas W., "Financial Intermediation and Delegated Monitoring," *Review of Economic Studies* 51 (1984), 393-414.

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

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

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

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

Goldsmith, Raymond, *Financial Structure and Development*, New Haven, Yale University Press, 1969.

Greene, William H., *Econometric Analysis*, 2nd Edition, Englewood Cliffs, NJ, Prentice Hall, 1990.

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

Grinblatt, Mark, and Sheridan Titman, *Financial Markets and Corporate Strategy*, Boston, Irwin/McGraw-Hill, 1998.

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

Hamilton, James D., *Time Series Analysis*, Princeton, NJ, Princeton University Press, 1994.

Hsiao, Cheng, *Analysis of Panel Data*, Cambridge, UK, Cambridge University Press, 1986.

Jayaratne, Jith, and Philip E. Strahan, "The Finance-Growth Nexus: Evidence from Bank Branch Deregulation," *Quarterly Journal of Economics* 111 (1996), 639-670.

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

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

King, Robert G., and Ross Levine, "Finance and Growth: Schumpeter Might Be Right," *Quarterly Journal of Economics* 108 (1993a), 717-737.

King, Robert G., and Ross Levine, "Finance, Entrepreneurship, and Growth: Theory and Evidence," *Journal of Monetary Economics* 32 (1993b), 513-542.

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

Levine, Ross, "Financial Development and Economic Growth: Views and Agenda," *Journal of Economic Literature* 35 (1997), 688-726.

Levine, Ross and Sara Zervos, "Stock Markets, Banks, and Economic Growth," *American Economic Review* 88 (1998), 537-558.

Lucas, Robert E. Jr., "On the Mechanics of Economic Development," *Journal of Monetary Economics* 22 (1988), 3-32.

Lucas, Robert E. Jr., "Why Doesn't Capital Flow from Rich to Poor Countries?" *American Economic Review* 80, Papers and Proceedings, (1990), 92-96.

Mookherjee, Dilip, and Ivan Png, "Optimal Auditing, Insurance, and Redistribution," *Quarterly Journal of Economics* 104 (1989), 399-415.

Nickell, Stephen, "Biases in Dynamic Models with Fixed Effects," *Econometrica* 49 (1981), 1417-1426.

Pagano, Marco, "Financial Markets and Growth: an Overview," *European Economic Review* 37 (1993), 613-622.

Rajan, Raghuram G. and Luigi Zingales, "Financial Dependence and Growth," *American Economic Review* 88 (1998), 559-586.

Romer, Paul M., "Increasing Returns and Long Run Growth," Journal of Political Economy 94 (1986), 1002-1037.

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

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

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

Sussman, Oren, "Economic Growth with Standardized Contracts," *European Economic Review*, forthcoming, (1999).

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

Warner, Jerold B., "Bankruptcy Costs: Some Evidence," *Journal of Finance* 32 (1977), 337-347.

Weiss, Lawrence A., "Bankruptcy Resolution: Direct Costs and Violation of Priority of Claims," *Journal of Financial Economics* 27 (1990), 285-314.

Zilibotti, Fabrizio, "Endogenous Growth and Intermediation in an Archipelago Economy," *Economic Journal* 104 (1994), 462-473.

Appendix A

Proof of Proposition 1:

The period t equilibrium conditions are equations (8), (9), and (11), namely:

(A.1) $$(1-p)w_{t+1}k_t\Phi(d_t) - BR = 0,$$

(A.2) $$pk_t\, l^d(w_{t+1}) + (1-p)k_t\Theta(d_t) - 1 = 0,$$

(A.3) $$p[a + \pi(w_{t+1})] - R - E - (1-p)w_{t+1}\Psi(d_t) = 0.$$

Hence, the equilibrium variables k_t, d_t, and w_{t+1} do not depend on any predetermined variables and the equilibrium is therefore time invariant.

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

(A.4) $$w = w(k,d) = (l^d)^{-1}\left[\frac{1}{pk} - \frac{1-p}{p}\Theta(d)\right].$$

Substituting (A.4) into equations (A.1) and (A.3) we derive two equilibrium conditions with the two unknowns k and d, describing the financial and the real sector:

(A.5) $$w(k,d)\Phi(d)k - \frac{B}{1-p}R = 0,$$

(A.6) $$pa + p\pi[w(k,d)] - R - E - (1-p)w(k,d)\Psi(d) = 0.$$

We next examine under what conditions these two equations determine a unique and stable equilibrium. Consider a limit case, where $1 - p$ tends to zero and B as well so that $B/(1-p)$ is constant. In this case w depends on k only as implied by (A.4). Hence, in this case (A.5) determines a downward sloping curve in the (k, d) plane and equation (A.6) determines an upward sloping curve. It can, therefore, be shown that in this case the equilibrium exists, it is unique and stable as well. Due to continuity, this is the case when $1 - p$ and B are sufficiently small. QED

Appendix B

In this appendix we demonstrate that the strength of the specialization effect depends on the level of output or wages. Consider the following example, where the production function is Cobb-Douglas:

$$g(l) = \frac{A}{\alpha} l^{\alpha}, \tag{B.1}$$

and the monitoring function is:

$$b(z) = b_0 + b_1 z, \tag{B.2}$$

where b_0, $b_1 \geq 0$. In the case of constant monitoring costs, namely $b_1 = 0$, we also assume that the setup banking costs B are zero, to rule out natural monopolies.

It can be shown that under the above specification, equation (16) becomes:

$$f = (1-p)\left[b_0 w + \frac{3}{2} b_1 (wd) \right], \tag{B.3}$$

where wd is determined by the following equation:

$$\frac{2b_1}{BR}(wd)^2 - \frac{b_1}{2} wd = b_0 w + \frac{p}{1-p} A^{\frac{1}{1-\alpha}} w^{\frac{-\alpha}{1-\alpha}}. \tag{B.4}$$

From equations (B.3) and (B.4) we draw the following conclusions:

1. If $b_1 = 0$, the wage effect dominates everywhere and ff is upward sloping.
2. If $b_0 = 0$, the specialization effect dominates everywhere and ff is downward sloping.
3. If b_0, $b_1 > 0$, the specialization effect dominates for low w, and the wage effect dominates for high w.

Tables

TABLE 1: SAMPLE MEANS, PANEL DATA 1982-1994

VARIABLE	SAMPLE MEAN	STD. DEV.
COST/GLL	5.85%	1.25%
DEPO/GLL	27.24%	8.01%
WRTOFF/GLL	1.13%	0.80%
GSP/CAPITA	$21,529	$3944

TABLE 2: OLS ESTIMATION, 1982-1994

Equation	α2	α3	α4	α5	α6	R^2	χ^2 (5%)	obs
(18)	0.713 (11.8)	0.129 (13.4)	-.018 (2.56)			.74	284 (3.84)	624
(19)	0.523 (14.0)	0.038 (3.02)	-.008 (2.45)	0.687 (12.1)		.90	0.867 (3.84)	576
(20)	0.527 (13.9)	0.035 (2.79)	-.008 (2.37)	0.681 (11.8)	-.001 (1.84)	.90	0.955 (3.84)	576

The reported coefficients are all estimated from the following general equation:

$$COST/GLL = \alpha 1 + \alpha 2 WRTOFF/GLL + \alpha 3 DEPO/GLL + \alpha 4 \log(GSP/CAPITA)$$
$$\alpha 5 lag(COST/GLL) + \alpha 6 \log(GLL/GSP) + e$$

Year and state dummies are included and jointly significant, but are not reported. Absolute values of t-stats are shown in parentheses using White heteroskedasticity-corrected standard errors, "obs" denotes the total number of observations, "χ^2" reports the Breusch-Godfrey chi-square test for serial correlation in the residuals (with the critical value for rejecting no serial correlation at the 5 percent error level shown in parentheses), and "R^2" is the adjusted R-squared statistic.

TABLE 3: GRANGER CAUSALITY, 1982-1994

Dependent Variable	COST -1	COST -2	COST -3	logGSP -1	logGSP -2	logGSP -3	R^2
COST	0.797 (10.56)	0.002 (0.02)	-.009 (0.11)	-.019 (2.76)	0.005 (0.42)	0.013 (1.45)	.88
logGSP	0.264 (1.27)	0.254 (0.92)	0.010 (0.05)	1.174 (29.08)	-.203 (3.19)	0.028 (0.62)	.98

The coefficient estimates are from the following two equations:

$$\text{logGSP} = \alpha 1 + \alpha 2 \text{ COST}(-1) + \alpha 3 \text{ COST}(-2) + \alpha 4 \text{ COST}(-3)$$
$$+ \alpha 5 \text{ logGSP}(-1) + \alpha 6 \text{ logGSP}(-2) + \alpha 7 \text{ logGSP}(-3) + e,$$

$$\text{COST} = \alpha 1 + \alpha 2 \text{ COST}(-1) + \alpha 3 \text{ COST}(-2) + \alpha 4 \text{ COST}(-3)$$
$$+ \alpha 5 \text{ logGSP}(-1) + \alpha 6 \text{ logGSP}(-2) + \alpha 7 \text{ logGSP}(-3) + e,$$

where COST = (COST/GLL) and logGSP = log(GSP/CAPITA). WRTOFF/GLL and DEPO/GLL are included and significant in each equation but are not reported (the results are robust to their exclusion). Absolute values of t-stats are shown in parentheses using White heteroskedasticity-robust standard errors, "R^2" is the adjusted R-squared statistic, the constant is significant but not reported, and the number of observations is 480.

TABLE 4: 2SLS ESTIMATION, 1982-1994

Instrument \ Equation	First Stage R^2	α2	α3	α4	α5	χ^2 test (critical value) (5%)
OLS	NA	0.523 (14.0)	0.038 (3.02)	-.008 (2.45)	0.687 (12.1)	NA
(IV 1)	.986	0.512 (13.6)	0.036 (2.78)	-.009 (3.18)	0.697 (12.2)	NA
(IV 2)	.829	0.480 (12.5)	0.014 (2.41)	-.003 (2.03)	0.776 (21.1)	NA
(IV 3)	.549	0.476 (3.72)	0.014 (6.32)	-.002 (2.46)	0.779 (2.48)	11.86 (22.4)
(IV 1a)	.987	0.514 (13.7)	0.036 (2.79)	-.009 (3.01)	0.691 (11.9)	0.01 (3.84)
(IV 2a)	.894	0.478 (12.4)	0.014 (2.38)	-.002 (1.96)	0.777 (20.9)	0.48 (3.84)
(IV 4)	NA	0.440 (7.89)	.003 (0.33)	-.011 (2.64)	0.759 (17.2)	0.20 (28.9)

The reported coefficient estimates are all estimated from the following equation:

$$\text{COST/GLL} = \alpha 1 + \alpha 2\ \text{WRTOFF/GLL} + \alpha 3\ \text{DEPO/GLL} + \alpha 4\ \text{GSP\_IV} + \alpha 5\ \text{lag(COST/GLL)} + e$$

where "GSP_IV" is the estimate of log(GSP/CAPITA) from the first stage regression:

$$\log(\text{GSP/CAPITA}) = c1 + \gamma \mathbf{Z} + \beta \mathbf{X} + u$$

where **Z** represents the three different sets of instruments and **X** is the vector of exogenous regressors. For (IV 1) **Z** equals lagged GSP/CAPITA, for (IV 2) **Z** equals GSP/CAPITA in 1982, and for (IV 3) **Z** equals demographic and regional variables from the early 1980's which can be viewed as exogeneous initial conditions for output. (IV 1a) and (IV 2a) are the same as (IV 1) and (IV 2), respectively, except their **Z** is augmented with the percent of state output from manufacturing. (IV 4) is (IV 3) with first stage regressions also for WRTOFF/GLL and DEPO/GLL, using the demographic and lagged **X** variables as instruments. The χ^2 test is the standard Hausman-type over-identifying test for rejection of the exogeneity of the instruments. It is computed as $T*N*R^2$ from the regression of the second-stage residuals on the first stage instruments.

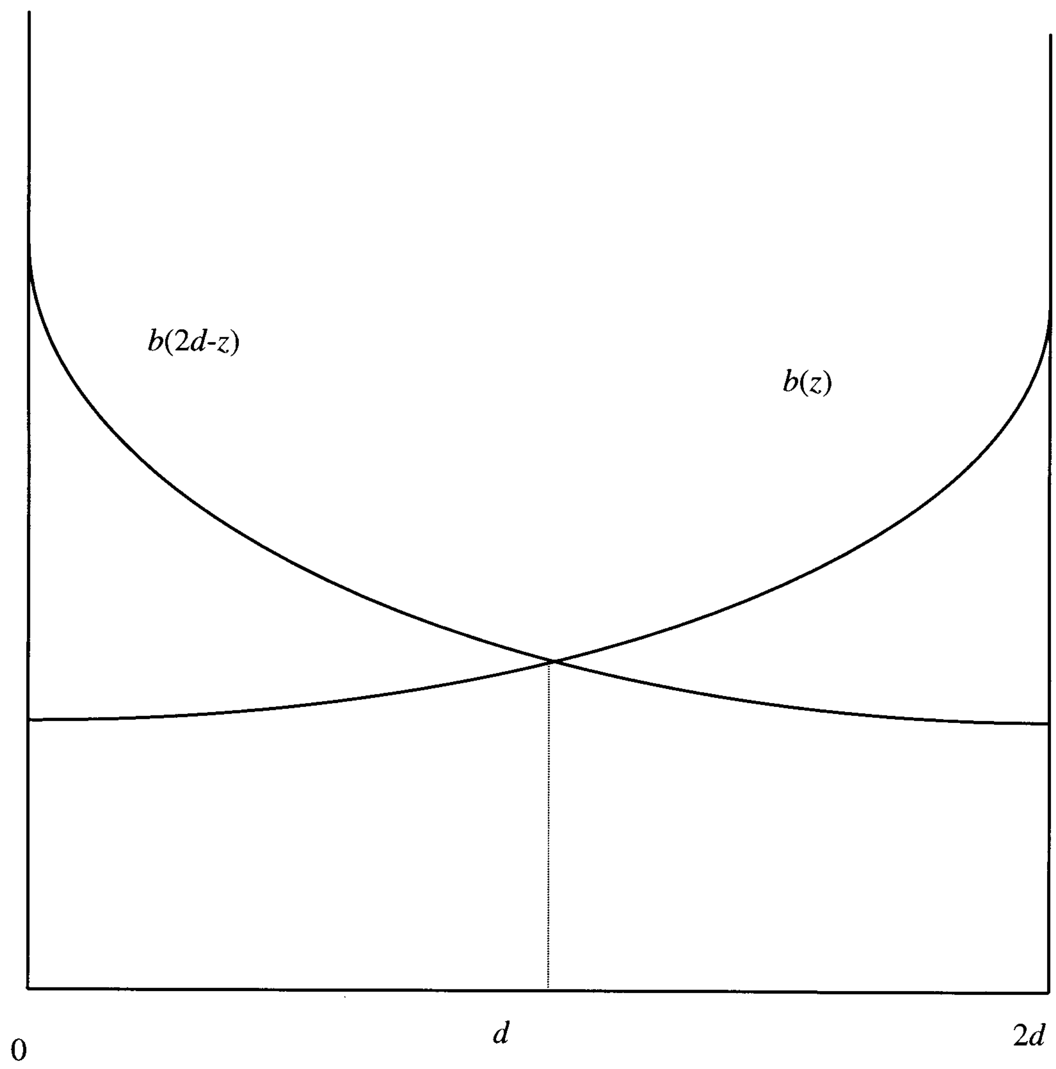

Figure 1

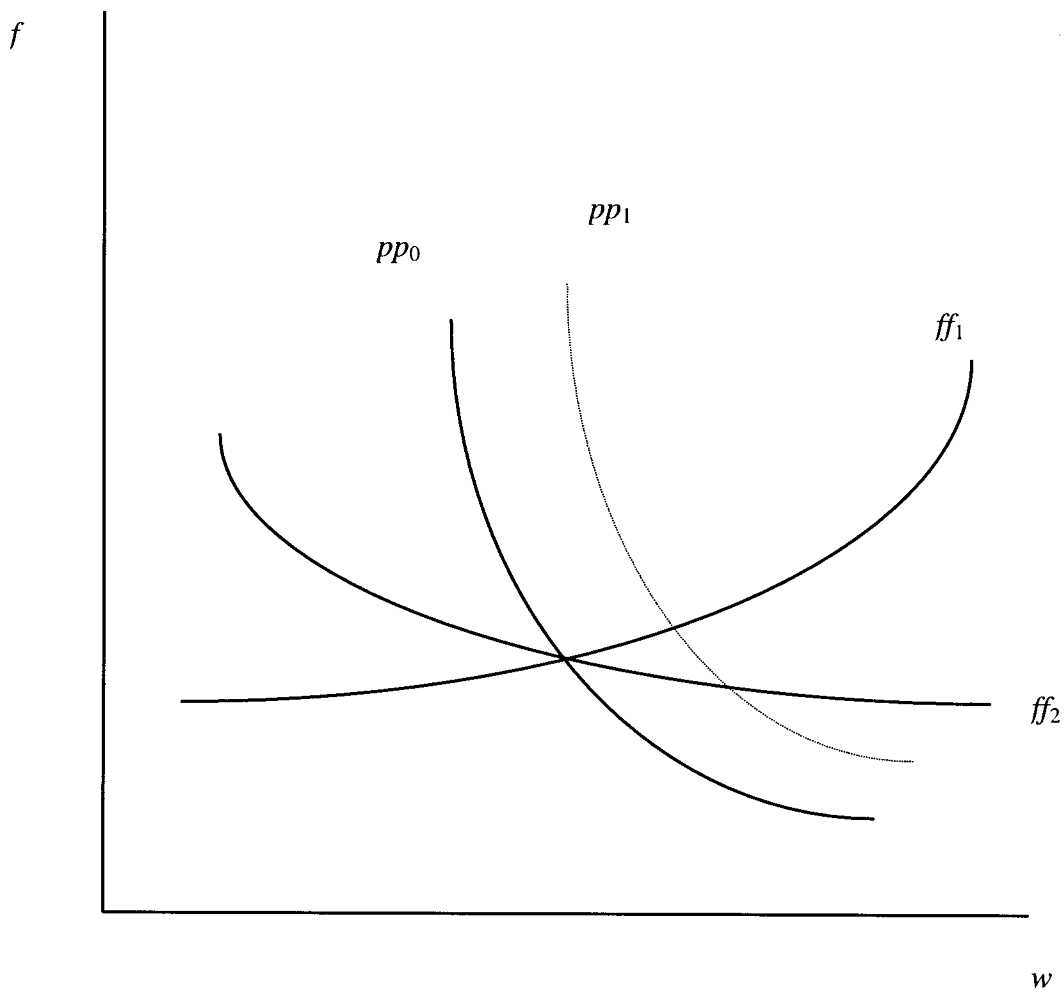

Figure 2

Figure 3. The Cost of Banking, 1985

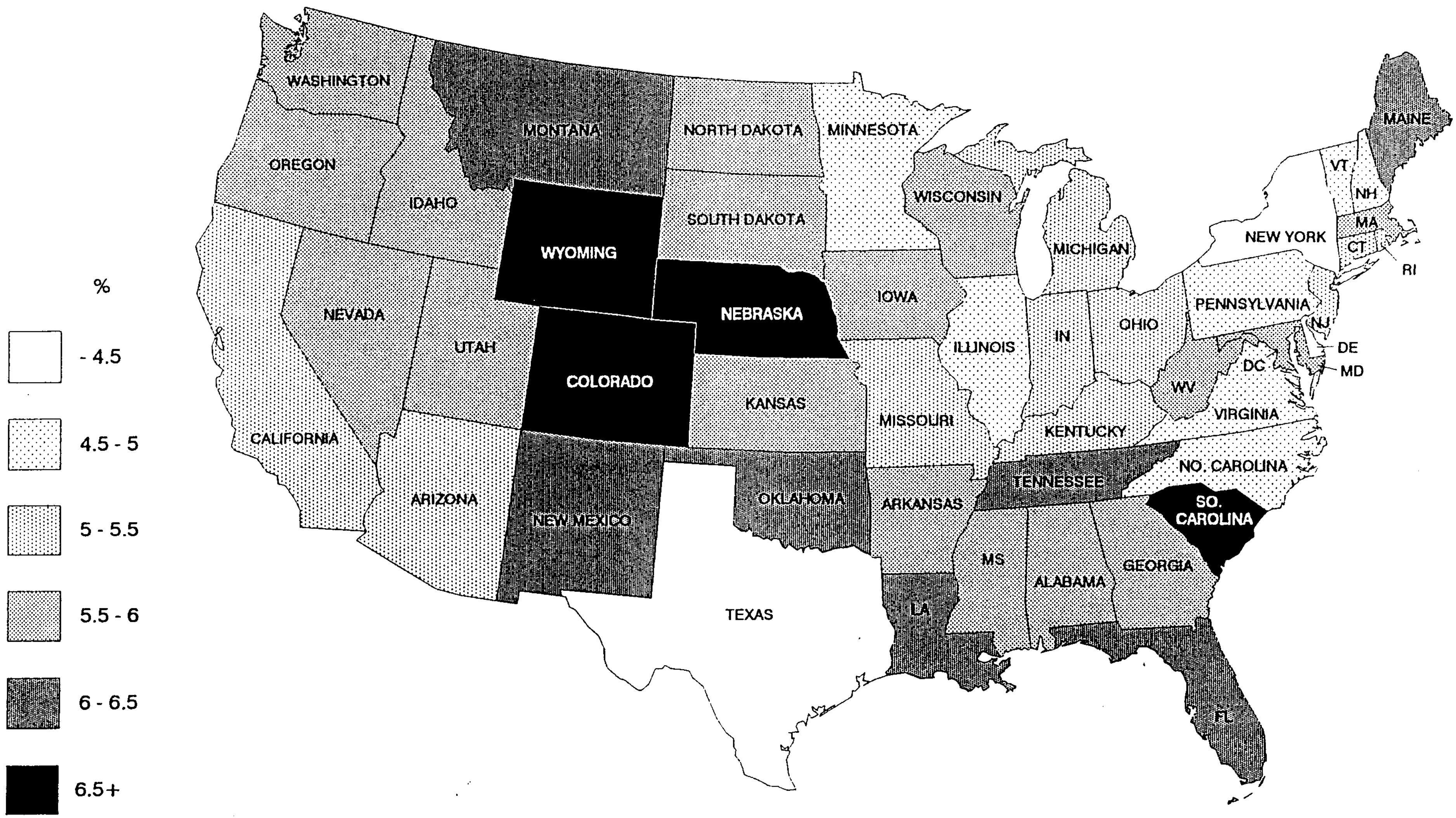

Figure Four: Year by year movements in cross-sectional means.