Economic Review

September/October 1992
Volume 77, Number 5

Federal Reserve Bank of Atlanta

In This Issue:

Economic Review

September/October 1992, Volume 77, Number 5

Economic Review

Federal Reserve Bank of Atlanta

The *Economic Review* of the Federal Reserve Bank of Atlanta presents analysis of economic and financial topics relevant to Federal Reserve policy. In a format accessible to the nonspecialist, the publication reflects the work of the Research Department. It is edited, designed, produced, and distributed through the Public Affairs Department.

Views expressed in the *Economic Review* are not necessarily those of this Bank or of the Federal Reserve System.

Free subscriptions and limited additional copies are available from the Public Affairs Department, Federal Reserve Bank of Atlanta, 104 Marietta Street, N.W., Atlanta, Georgia 30303-2713 (404/521-8020). Change-of-address notices and subscription cancellations should be sent directly to the Public Affairs Department. Please include the current mailing label as well as any new information. ISSN 0732-1813

Contents

September/October 1992, Volume 77, Number 5

Human Capital Investment And Economic Growth: New Routes in Theory Address Old Questions

Ellis W. Tallman and Ping Wang

The question of how economies grow has been a central topic of research for economists since the time of Adam Smith. The formal analysis of growth focuses on long-run economic progress, and since the mid-1950s the dominant theoretical framework for inquiry into economic development, the model proposed by Robert M. Solow (1956) and T.W. Swan (1956), has provided a paradigm for analyzing growth that formalizes how inputs of physical capital and raw labor combine to create real (inflation-adjusted) output. In related empirical work Edward F. Denison (1962) employed statistical methods in a procedure referred to as growth accounting to investigate the sources of income growth. Denison focused on the growth rates of physical capital and raw labor to determine how much income growth can be explained by the growth in inputs.

Unfortunately, in applications of the Solow-Swan model as well as growth accounting the empirical results show that the input growth rates fail to explain most of the variation in output. In general, the unexplained portion of output growth has been attributed to the area of technological progress, where measures of inputs do not capture improvements in output creation such as new methods of production and innovations in transportation.

Theodore W. Schultz (1961) attributed the fact that the growth in output exceeds that of the measured inputs—the components of production, physically reproducible capital and raw labor worker-hours—to investment in human capital. His analysis focused on the concept that individuals invest in learning skills, gaining knowledge, and otherwise enhancing their physical or mental abilities. According to Schultz, improvements embodied in inputs, such as technological changes, are likely the products of human capital in

Tallman is an economist in the macropolicy section of the Atlanta Fed's research department. Wang is an assistant professor in the Department of Economics of The Pennsylvania State University and visiting senior economist at the Federal Reserve Bank of Dallas.

action, especially in the United States, where much leisure time is spent enhancing skills and gaining knowledge. His intuitive argument implies that technological change, often viewed as an exogenous or external factor affecting the economy, may be explainable within a model embodying the human capital investment decision. While Schultz sought to address existing research on economic growth and development, he also related his comments to studies examining the returns to education, which now make up a voluminous body of literature but were new discoveries at the time of his writing.

The literature on human capital and growth highlights the concept of human capital as a mechanism to advance technology, improve productivity, and generate growth. This article reviews that research and elaborates Schultz's assertion appending human capital to the Solow-Swan growth model. A discussion of the relationship between education and lifetime earnings presents relevant empirical evidence. The article then synthesizes the recently developed endogenous growth literature—that is, the studies examining forces within the economy that may generate growth. The analysis considers the role of human capital in enhancing output growth, emphasizing how the rational decisions of individuals choosing to invest in education may make a difference. Both theoretical and empirical deficiencies in the existing studies are summarized. In conclusion, the article points out policy implications of the findings concerning human capital and its relationship to output growth.

Technological change, often viewed as an external factor affecting the economy, may be explainable within a model embodying the human capital investment decision.

Observing that international trade and competition have drastically reduced earnings potential for lesser-skilled workers, policy-oriented research has proposed that future job growth will be in technical fields, requiring workers to have advanced education levels. The crises in U.S. cities among the unskilled unemployed and in education systems illustrate the need for investments in support of human capital formation. It is appropriate that the topic of human capital has received attention not only from U.S. policymakers and the economics profession but also from the business world (see, for example, Bruce Nussbaum et al. 1988 and Leonard Silk 1992).

The Solow-Swan Growth Model and Growth Accounting

The Solow-Swan model in many ways revolutionized the theory as well as the measurement of economic growth. Its impact has been widespread across various economic disciplines.[1] The characteristics described here will paint at least a superficial picture of this influential model.

The Solow-Swan growth model begins with a basic assumption that output (Y) can be produced using combinations of physical capital (K) (including reproducible machines, equipment, and buildings) and labor (L) in variable proportions. The model assumes constant returns to scale (*CRTS*), implying that doubling the amount of each input in combination will double the output. However, if either input is held fixed and the other is doubled, output will increase by less than double.[2] In addition to using labor and capital as the only inputs, the economy's production function is subject to a technological factor (A) in the standard Solow-Swan model representation.[3] The production function for the economy, often referred to as the aggregate production function, is given below:

$$Y = F(K, AL), \tag{1}$$

where the production function, F, determines how much output, Y, can be made by combining inputs of capital and labor, K and L, in the production process. The factor A represents the labor-augmenting technological advancement, which is exogenous. The labor-augmenting character of the technological factor means that it is like simply adding more labor to the production function. This specification of the factor is referred to as Harrod-neutral technological change.

The model's steady-state relationships—in which all variables grow at constant rates—highlight key implications that have been a source of criticism. One important equation can be derived directly from the definition of the growth rate of capital and the

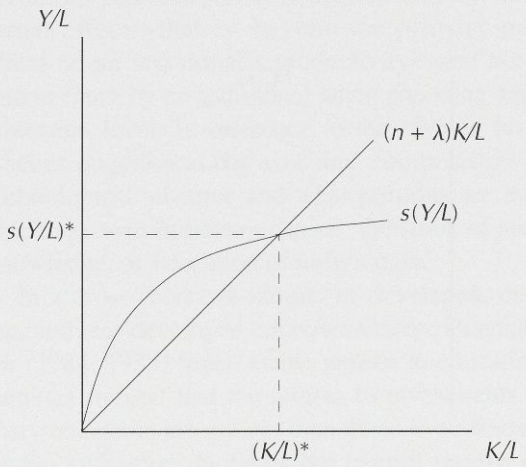
equation for aggregate investment/savings. The steady-state equation is below:

$$s(Y/L) = (n + \lambda)K/L, \qquad (2)$$

where s is the savings rate, n is the rate of population growth, and λ is the rate of technological advancement. With constant returns to scale, per capita output (Y/L) is a function of the capital-labor ratio (K/L) alone. Thus the equation can be used to determine the steady-state accumulation of physical capital in per capita units (also see Chart 1). In the equation, two exogenous processes (n, λ) determine the economy's growth rate. The savings rate determines the equilibrium output/labor and capital/labor ratios, but increased savings cannot alter the rate of economic growth in the model. In addition, the population growth rate cannot alter the per capita growth in output. The only way to achieve higher per capita output growth is to have a greater rate of technological advancement, something presumed to be outside the control of the agents in the economy.

To explore the role of technological change for explaining growth in the United States, Solow (1957) employed a simple empirical analogue to the growth model. In it, he estimated a series he referred to as "technological change" (often called the "Solow residual") that reflects the concept of an external force generating economic growth. The notable finding was that almost 90 percent of the growth in the U.S. economy from 1909 to 1949 could be attributed to the technological factor.

Chart 1
Steady-State Capital Accumulation In the Solow-Swan Model



Denison (1962, 1974) provided alternative estimation of economic growth, referred to as growth accounting. The empirical method begins with the basic assumption that output depends on numerous determinants (inputs) and that changes in these inputs cause output changes. How much growth a particular source contributes depends upon its importance to output growth as shown in the amount that output growth changed. The statistical methods attempt to measure the sources of economic growth using data on relevant inputs.[4] The typical growth accounting equation is

$$\Delta Y/Y = \alpha(\Delta L/L + \Delta A/A) + (1 - \alpha)\Delta K/K, \qquad (3)$$

where $\Delta A/A$ represents the growth rate of labor-augmenting technological change, α is the labor income share (that is, the return to labor from the production of output), and $(1 - \alpha)$ is the capital income share.[5] Denison measured the growth in labor hours and capital stock and compared these variables with the growth in output. In his 1962 study Denison found that in the United States over the 1909-57 period real output, employed worker hours, and capital stock grew an annual average rate of 2.9 percent, 1.4 percent, and 2.4 percent per capita, respectively.[6] That part of output growth not accounted for by input growth, the residual measure that he interpreted as "advances in the state of knowledge," is comparable to the rate of technical progress in Solow (1957). Many researchers have found the significant proportion of output attributed to an external (exogenous) process of knowledge or technological advancement an obvious flaw in the theory and its application. The model appeared to leave an inordinate amount of variation in economic growth unexplained.

*E*ducation and Human Capital Investment

Schultz (1961) emphasized a linkage between earnings and education, suggesting that investment in human capital accounts for most of the observed rise in real earnings. In addition, he proposed that the value of human capital may be as big or bigger than the value of the tangible physical capital stock, with a rate of growth in excess of capital growth.[7] Schultz also acknowledged that improvements in factor input quality generally affect productivity. For example, suspension bridge construction methods illustrate how input quality has improved over time, increasing productivity. The

construction of the Brooklyn Bridge took fourteen years in the 1880s; during the 1960s the Verrazano-Narrows Bridge, more than twice the span of the Brooklyn Bridge, was built in five years. Such improvements in productivity may result from advances in physical technology (new capital goods), the stock of knowledge, or human capital embodied in individual workers. The matter is a key issue in research.

Although the idea of investment in human capital may seem intangible and difficult to quantify for scientific analysis, measuring the individual's education levels has been a successful method. Gary Becker (1975), for instance, investigated the returns to education in the United States. The theory analyzes the educational choice of consumers, using a model in which individuals choose the level of education that they desire on the basis of expected returns to their investment of time, effort, and expense. Becker explicitly linked the education level of an individual to his or her productivity as a worker, implying that those workers with higher education are more productive and therefore receive higher wages. The human capital investment function is as follows:

$$HCINV = G(R, T, B, H), \tag{4}$$

where $HCINV$ is the rate of investment in human capital, G is the human capital investment function, R is the input of other resources (capital and labor), T is the input of time toward education, B is the physical and mental powers of the individual, and H is the input of human capital. Additional amounts of any of these inputs result in increased production of human capital—that is, human capital investment is a positive function of all the inputs.

In 1976 Sherwin Rosen published a breakthrough article that introduced dynamics to the theory of life earnings relying on human capital investment. The lifetime earnings function is below:

$$W(t) = \int_t^N E(H, \dot{H}, s)e^{-r(s-t)}ds, \tag{5}$$

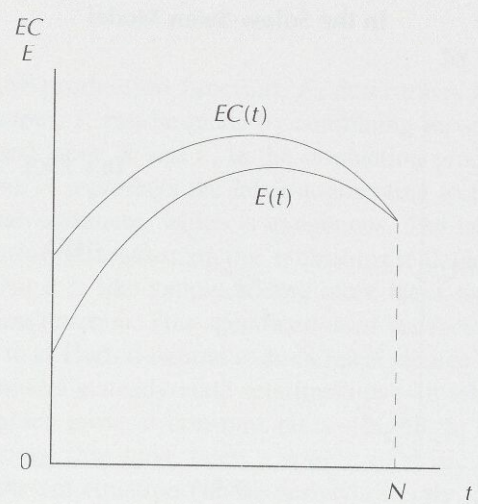
where $W(t)$ represents lifetime earnings, N is the predetermined length of the work life, E is the current earnings function, s measures time, and r is the discount factor.

The model takes into account future as well as current earnings, discounting future earnings and presenting a lifetime earnings function. In this framework, current earnings are positively related to the accumulated human capital stock (H) but are negatively related to increased current investment in human capital ($\dot{H}$) because it takes time and resources away from current work effort and wages. Clearly, there is a trade-off between current earnings and increased future productivity by further investment in human capital today, and the model takes account of these factors and determines an optimal level of human capital investment.

Chart 2 reflects the earnings capacity (EC) and actual earnings (E) over an individual's lifetime. Earnings capacity exceeds actual earnings, and the difference between them reflects human capital investment. The chart illustrates the tendency of the lifetime pattern of earnings to peak during middle age and to show the greatest difference between earnings and earnings capacity in the early years. Human capital investment during the early stages of a lifetime has a long time to generate returns. An individual who invested less in human capital would have flatter curves on this chart.

Rosen's work is important not only for adding the temporal dimension to the analysis of returns from education but also because it elaborates on the various ways that education enhances human capital. Certain processes that take place on the job—for example, learning by doing and specific job training—add to human capital by increasing worker productivity. Rosen's theory emphasizes that education both improves an individual's knowledge once and for all and strengthens the individual's capacity to learn on the job, thereby increasing the worker's speed of human capital accumulation. This insight implies that the labor productivity enhancement from investment in human capital may exceed that recognized by Becker (1975).

Chart 2
Earnings Capacity versus Lifetime Earnings



*E*mpirical Evidence on Returns to Education

The evidence on returns to investment in education is voluminous, and the selective sample offered here attempts only to raise some issues that relate, directly and indirectly, to inquiries on economic growth. The work of T. Paul Schultz (1988), for example, raises additional questions about our understanding of the sources of growth, suggesting that although it may be comfortably assumed that the changing quality of capital and labor are likely sources of growth, accounting for such sources explicitly remains unresolved. Schultz presented a useful survey of the literature on the returns to education as well as the link between education and growth and concluded that "the record of sustained modern economic growth in real per capita income cannot be accounted for by the accumulation of conventional units of physical capital or by the increased application of hours of labor per capita" (1988, 544).

Becker's (1975) empirical work supports his premise that more education is correlated with higher earnings. At the same time, his interpretations of several results hint at more general issues that have become central in the recent endogenous growth literature. For example, he suggested that gains from college education are not fully quantified by earnings analysis because college graduates are only partially compensated for their effect on the development and spread of knowledge. He observed that the accumulation of education (by individuals) may be measured and conceived of as separate from the growth in knowledge (or technology), although increases in both education and knowledge improve productivity. Becker's interpretations perceive part of education's effects as an external effect—that is, beyond the primary positive effect on an individual's productivity—and not taken into account by an individual when choosing a desired education level. In reference to the "Solow residual" Becker questioned the size and composition of the technological advance and what portions are attributable to growth in human capital, the stock of society's knowledge, or education of individuals.[8]

In a cross-sectional sample of developed, developing, and less-developed countries George Psacharopoulos (1984, 1985) found stable returns to education. His findings suggest that the returns to investments in primary education are greater than those to secondary and higher education. Notably, the overall returns profile declines as the education level increases. The results, however, cannot capture the potential external effects of higher education implied by Becker.[9]

Richard A. Easterlin (1981) examined a cross-section of countries and found a link between widespread public education and economic growth. In agreement with Becker's ideas on the external effects of education, Easterlin hypothesized that modern economic growth relies on the diffusion and the advancement of knowledge. He viewed the spread of mass education, separate from the growth of science and technology, as a key to economic development, offering upward mobility to a wider segment of the world population. His work suggests that a populace characterized by at least basic education is a precondition for economic growth and that widespread expansion of schooling reflects a voluntary move on the part of informed governments toward economic growth through education.

One of the more fundamental external effects from human capital investment lies in the advancement of knowledge and the development of new applications of knowledge.

Dale W. Jorgenson and Barbara M. Fraumeni (1991) presented measurement of investment in human capital from a perspective different from previous studies: they measured the investment in terms of the income produced by human capital rather than using the more typical method of examining the outlays for education.[10] Their study found that the overwhelming portion of economic growth in the United States is based on investment in both human and physical capital. Estimates show that in the United States investment in education dwarfs other kinds of investment. Jorgenson and Fraumeni also noted that their estimate of the stock of human wealth, derived from the measures of educational output, is ten times greater than previous estimates (see John W. Kendrick 1973).[11]

As Schultz (1988) pointed out, the existing evidence on education and returns suffers from some

measurement problems. First, the studies overlook immeasurable components like effort and innate ability. Second, although there is a correlation between education and income, the nature of the underlying link between them is uncertain. For example, the rich are generally well paid and well educated. Does higher income result from better education or from nonhuman (financial) wealth?

An additional concern is whether, in certain environments, there may be overinvestment in education. Psacharopoulos (1985), for instance, argued that certain underdeveloped nations have placed too much emphasis on higher education without enough attention having been placed first on primary education; the profile of estimated returns suggests that the composition of education expenditures should be skewed toward widespread elementary education first.[12] However, the assumption that there is overinvestment in education may overlook the potential external effects of human capital (which will be further examined below) so that the returns resulting from additional human capital are not linked to investments made in human capital accumulation.

*H*uman Capital Accumulation And Output Growth

A relatively new direction of study has arisen to explain economic growth and development without appealing to an exogenous source of technological advance as the main source of economic growth. The models offer a way to identify the role of human capital in enhancing output growth, to emphasize individuals' decisions to invest in acquiring skills, and to rationalize how these actions allow the economy to grow endogenously—that is, as a result of the actions of individuals represented in the model. In this framework human capital accumulation provides the engine of growth by achieving the technological advance that previous models assumed to be exogenous. Thus, the endogenous growth theories attempt not only to identify the main sources of technological change (such as endogenous human capital accumulation) but also to design models in which economic incentives (such as greater returns to higher levels of education) explain what drives economic advancement.

The aggregate production function that incorporates the endogenous growth feature is

$$Y = AF(K, HL), \tag{6}$$

where the disembodied technological factor, A, is outside the function and human capital, H, is labor-augmenting technical change. The function is assumed to display constant returns to scale in two reproducible capitals—physical, K, and human, H—in contrast to diminishing returns in the Solow-Swan model.

One of the main issues of contrast between the Solow-Swan and the endogenous growth models concerns the predicted growth rate of output per effective unit of input in the steady state, or the long run. The Solow-Swan growth model predicts a zero growth rate of output per effective unit input because output growth is entirely determined by exogenous factors like the population growth rate (affecting labor input) and the labor-augmenting technology shock. The long-run capital and effective labor growth rates are the same as the exogenous rate of growth—the sum of the growth rates of population and the technology shock. If a Solow-Swan growth model is examined without the exogenous rate of technological advancement, the rate of per capita growth in the model economy is zero. On the other hand, for endogenous growth models the growth rate of output per capita is a positive constant because the advancement in technology results from the choice of individuals to invest in human capital. As human capital accumulates, technology improves. Rather than exogenous factors determining growth, the technological advances that enhance productivity are attained endogenously.

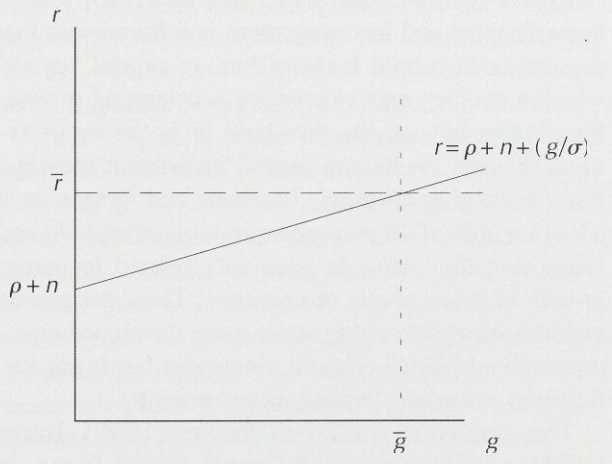
The rate-of-returns equation from an endogenous growth model is useful for distinguishing between the extended, optimizing Solow-Swan and endogenous growth models (also see Chart 3):[13]

$$r = MPK = \rho + n + (g/\sigma), \tag{7}$$

where MPK is the marginal product of capital (which equals the real rate of return, r), ρ is the rate of time preference (which measures a consumer's preference for present rather than future consumption), n is the rate of population growth, g is the rate of (endogenous) economic growth, and σ is the intertemporal elasticity of consumption substitution (which measures the willingness of consumers to substitute between current and future consumption).

The characterization of the marginal product of capital is a key difference between the two theories. In both models the rate of population growth, the time preference rate, and the intertemporal elasticity of substitution are generally assumed to be exogenous. To determine the equality, either the rate of

Chart 3
Exogenous versus Endogenous Growth



economic growth or the rate of return to capital must be determined a priori. The key difference between the models involves the choice of which rate will be considered endogenous (determined by the equation).

The Solow-Swan model assumes diminishing returns to capital, implying a marginal product of capital that falls as additional capital accumulates, holding other inputs fixed. As a result, the rate of return to capital is endogenous, depending on the level of the capital-labor ratio. However, as described above, the rate of growth in the model economy is constant and externally (exogenously) determined. In contrast, endogenous growth models imply that the marginal product of capital is constant—that is, the rate of return on capital, r, is fixed and determined solely by the production technology. This restriction is implied by the assumption of constant returns to scale in the reproducible factors so that additional amounts of capital do not reduce the rate of return. These assumptions allow the model to determine the growth rate (g) endogenously. Thus, the Solow-Swan model allows the marginal product of capital to vary but fixes the rate of economic growth while endogenous growth models fix the marginal product of capital but allow the rate of economic growth to be endogenous.

In Chart 3 the upward-sloping line is the Keynes-Ramsey formula (equation [7]) for the real interest rate. The vertical line at $g = \bar{g}$ reflects how the Solow-Swan model fixes the growth rate but can determine the rate of return to capital because of diminishing returns to capital. In contrast, the horizontal line at $r = \bar{r}$ indicates how the endogenous growth models fix the rate of return based on constant returns to reproducible capitals but can determine the (endogenous) growth rate.

The amount of research in this field has grown rapidly. While a comprehensive survey of the literature is beyond the scope of this article, it should be useful to present several of the main ideas that characterize the role of human capital in an aggregate production function and to provide some empirical evidence that relates to the predictions of theory.[14]

Robert E. Lucas (1988) provided a clear-cut linkage between the aggregate production models of growth (Solow-Swan) and the idea that human capital levels directly affect output. The model suggests that human capital accumulation is the main driving force of economic growth. Moreover, Lucas's work emphasizes an external effect of human capital—that the average level of human capital can magnify the impact of individual human capital and lead to greater output. For example, the concentrated, collaborative effort of great scientists on the Manhattan Project produced the atomic bomb perhaps more quickly than if they had not benefited from the simultaneous group effort. The individual deciding to invest in human capital fails to account for this external effect of human capital, though, basing his or her choices solely on the perceived private returns. Thus, one of the main implications of Lucas's findings is that human capital investment is likely to be below the socially optimal level unless there is market intervention in the form of a subsidy for accumulating additional human capital.

Unlike Lucas (1988), Paul Romer (1990) envisioned a model in which human capital is an essential source of economic growth but human capital levels have no external effect. Human capital has two different definitions—cognitive skills tied to individuals and aggregate knowledge not tied to individuals. Cognitive skills are often measured by variables like individuals' years of education, which cannot grow perpetually and are perceived as unlikely to explain the sustained growth of per capita output. However, the stock of human knowledge (scientific understanding) may grow without bound, and the key to growth and development is its exploitation. In Romer's model, effective use of human knowledge leads to developing new capital goods that are more productive than previous versions. Old capital goods may still be useful, but the new goods are more efficient. Consider, for example, the personal computer. Five years ago, the 80286-based PC was state-of-the-art; it remains useful now, but it is far less productive than either the 80386- or

80486-based machines. Thus, improved machines essentially embody the human capital of advanced knowledge.[15]

Nancy Stokey (1991) brought together two key points from the Lucas and Romer models. The Stokey model includes the external effect of education suggested by Lucas but also separates the stock of human knowledge from individuals as in Romer (1990). The external effect is seen as coming about from individuals' investment in human capital. Human capital accumulation is a conscious decision of an individual to invest and enter school, but as individuals accumulate human capital the aggregate stock of knowledge increases as well. Additional human capital improves labor quality, increases the stock of knowledge, and directly affects the product pattern, leading to the development of new and higher-quality goods.

The Lucas, Romer, and Stokey models present numerous insights into the characterization of human capital as the engine of growth in an endogenous growth model. The key aspects of each study revolve around the treatment of human capital in an aggregate production function and the question of how human capital enhances output. The empirical implications of the models remain somewhat limited, but the empirical work is enough to give a general impression of the evidence and the direction research is likely to take.[16]

*E*mpirical Evidence on Human Capital and Growth

Romer (1990) investigated whether the literacy rate in 1960, a proxy measure of the initial level of human capital, affects the growth experience of a cross-section of countries in the subsequent twenty-five years. The results suggest that the literacy rate fails to affect output significantly once the rate of investment is included in the analysis. On the other hand, Romer found that the initial level of human capital (and the change in literacy) does have a significant effect on the rate of investment, providing an indirect link between human capital and output growth through physical capital investment.

Robert J. Barro (1991) used enrollment rates in primary and secondary schools as human capital proxies in a cross-country empirical study. He found significant positive effects for primary and secondary enrollment rates for 1960 on output growth averaged over the period from 1960 to 1985. It should be noted, however, that enrollment rates for 1950 and 1970 are not significant for explaining average output growth, suggesting that the positive results based on 1960 enrollments may be not be robust.

Costas Azariadis and Allen Drazen (1990) related human capital and income growth in a framework that presents a threshold level of human capital beyond which a country may experience accelerating growth. Essentially, beyond the threshold level the social returns to scale for human capital investment increase. For empirical application, Azariadis and Drazen used a literacy rate of 40 percent as an initial threshold and found that this proxy is positively related to output growth in their sample of countries. Thus, the private yield on education is higher in more developed countries, and additional education precedes but is not sufficient to cause accelerating output growth.[17]

The empirical studies of Romer (1990), Barro (1991), and Azariadis and Drazen (1990) found, in varying degrees, some support for the idea that human capital has significant explanatory power for output growth. Their perspective is that human capital provides for endogenous growth. In contrast, N. Gregory Mankiw, David Romer, and David N. Weil (1992) designed human capital accumulation as an exogenous process rather than a function of individual decisions. Their evidence supports a role for human capital, but it is contrary to the idea that human capital has external effects and to the framework of perpetual growth. Using a transformation of secondary school enrollment as the human capital proxy, Mankiw, Romer, and Weil show that the role of human capital in the aggregate production function is consistent with diminishing returns to scale in all reproducible factors.[18] Unlike in the endogenous growth models, their results imply that perpetual output growth cannot emerge as a result of physical and human capital accumulation. They argue that the Solow-Swan growth model is preferable for analyzing economic growth.

While suggestive, the existing empirical results do not offer conclusive evidence in support of human capital having a significant role in economic growth. Ross Levine and David Renelt (1992) examined the robustness of correlations between long-run growth rates and policy variables (several of which are human capital proxies) found in cross-country empirical studies. Their findings indicate that regression results that capture a positive relationship between human capital and growth are not robust to the inclusion of other relevant variables.[19] As a consequence, they suggest a reasonable degree of skepticism about inferences from empirical studies linking human capital and growth.

In another recent study, Jess Benhabib and Mark M. Speigel (1991) estimated an aggregate production function specification of growth by constructing physical and human capital series for a selection of countries. Consistent with Romer (1990), their results suggest that human capital does not enter significantly into the explanation of aggregate output growth but remains important in explaining capital accumulation.

These empirical studies embody a number of issues to be addressed by future research. Measurement problems in particular are commonly pointed to as a criticism of these initial studies. One concern is that the studies examine cross-sections of data, and the data—how they were collected, how reliable the numbers are, and so forth—may not be consistent across countries. In addition, the human capital proxies are necessarily crude.[20] For instance, although the literacy rate may be a fairly consistently measured variable, it may only tangentially measure the human capital concept of interest (that is, a measure of knowledge or achievement, a more advanced human capital variable). Further, enrollment rates are flow variables that measure the proportion of the population attending educational institutions but do not necessarily pick up relevant movements in the stock of human capital or knowledge. Another limitation is that simply comparing average output growth rates with initial levels of some human capital proxy fails to examine some of the interesting dynamics that take place over time between human capital and growth.

In an alternative empirical strategy, Ellis W. Tallman and Ping Wang (1992) have focused on the growth experience of an individual country, Taiwan, over time to examine the effect of human capital on output. This approach has several advantages. First, there is a more appropriate measure of human capital that focuses on achievement levels—that is, an aggregate human capital measure created by weighting levels of education completed.[21] The human capital measure is used in combination with the raw labor measure to form an effective labor input. Estimating an aggregate production function yields the finding that effective labor added to measured physical capital directly affects output growth. In addition, it is shown that the income shares from most estimates are consistent with constant returns to scale and are robust to adding variables that are typically correlated to output in the output regressions. The research complements cross-country studies that attempt to find more general human capital/growth relationships.

Even among endogenous growth theories there are differences on a number of issues. Distinctions between human and physical capital are not made entirely clear, for instance. In Romer, human capital accumulation occurs, but it is new capital goods that provide the key input toward output growth. If human capital in the form of knowledge represents potential capital but only becomes a driving force of growth when it is embodied in physical capital, which of these factors is the true driving force? Another question centers on whether human capital depreciates. Capital embodied in individuals likely depreciates, but it is unlikely that the stock of knowledge declines in value (although it might be subject to shocks, such as the complete destruction of the Mayan libraries).

A further shortcoming is that the predictions of both the Solow-Swan and endogenous growth models appear somewhat unrealistic regarding observed behavior of economies. Instead of either zero or constant positive growth rates, there are variable growth rates of output per effective unit input depending on a country's stage of development. Less-developed countries appear to experience a low growth rate until a crucial (perhaps human capital level) threshold is passed. Newly industrialized countries may have high, sustained growth rates whereas developed countries may experience slower output growth.

It is also true that the two model frameworks appear difficult to disentangle empirically despite some obvious contrasts. The Solow-Swan model predicts that output levels across countries possessing the same parameters of technology and preferences will converge over time. According to this assumption, poor countries should grow faster than rich countries. In general, however, endogenous growth models offer no such prediction.[22] Empirical tests of this "convergence hypothesis" have not reached a consensus. The results of Mankiw, Romer, and Weil (1992) and Barro and Xavier Sala-i-Martin (1992) show evidence consistent with convergence; Barro (1991) finds evidence of income nonconvergence.[23] Unfortunately, the evidence supporting convergence suffers from a problem in the estimation methods that renders the empirical results equivocal.[24] To make things more complicated, some endogenous growth models imply convergence, and some extended Solow-Swan models predict nonconvergence.

It appears that in the literature on endogenous growth models theory may be ahead of the measurement. Applications of the endogenous growth literature are only beginning to confirm a significant relationship between human capital and growth (but not all implications from the models are directly testable). Improvements in formulating human capital

measures may establish a stronger link between human capital and growth. In addition, more intensive examinations of specific case studies may be a promising path toward understanding the role of human capital in economic development.

Conclusion

Numerous studies in the economics literature propose that through a variety of roles human capital provides increased earnings for individuals and generates economic growth. Although the empirical evidence in support of a link between education measures and output growth is equivocal, all studies suggest a beneficial effect of education on output or individual earnings.[25]

Few, if any, would argue that an economy can have too much human capital or too much education.[26] Increased demands for technically skilled workers make it clear that improved human capital is essential for economic progress. The economics literature offers an overview of the particular characteristics of human capital that may be important for growth, some of which may require activist policies to achieve the most desirable outcomes.

First, the diffusion of education may be essential to the spread of modern economic growth. Widespread public education at at least a basic level sets the stage for development, as in the idea of a threshold, discussed above (Azariadis and Drazen 1990). Such a policy is likely to be more appropriate for an economy in the early stages of development, when primary education should be a first priority. Educational improvements can be expected to allow a higher level of future growth because workers would be better able to operate and exploit modern, sophisticated physical capital.

In developed nations with widespread education systems, like the United States, the emphasis should be on the quality of education at the primary as well as higher levels. Greater investment in human and physical capital appears necessary for the United States to regain sustained economic growth. Recognizing problems in urban public schools may help galvanize a focus on the long-term economic benefits to be realized by improving the quality of our school systems nationwide. Unfortunately, because human capital involves a long gestation period and is costly considering the time value of the investment, no immediate solutions are at hand for the United States. With substantial investment and considerable patience we may at least anticipate future returns.

An important point to keep in mind is that individuals do not take into account the positive external effects of human capital. The overriding implication is that individuals are likely to underinvest in education and, in terms of enhancing growth, there may need to be subsidizing policies that encourage the accumulation of human capital. Such policies may help explain the rapid development of newly industrialized countries like Taiwan. More specifically, one of the more fundamental external returns to human capital investment lies in fundamental research in basic sciences. The advancement of knowledge and the development of new applications of knowledge—technological advancements—provide for future economic growth, and in the long run the United States would benefit from directing individuals into the fields of scientific research and engineering.

It is generally agreed that there are potential gains from greater emphasis on higher education, which improves learning efficiency on the job and yields significant positive external effects. This improvement in on-the-job learning is also important for promoting perpetual economic growth, adding significantly to individual human capital stock as well as to the stock of society's knowledge that may improve the quality of life.

Notes

1. Notably, in the analysis of real business cycles Prescott (1986) employs Solow's (1957) method for accounting for technological change. The Nobel Prize given to Solow indicates the significance of his work.
2. Economists refer to this characteristic of a production function as displaying the "diminishing marginal product" of the single factor.
3. The factor grows over time but may be subject to random shocks. This factor, proposed as an exogenous (that is, determined outside of the model) rate of technical change, has been interpreted as the rate of knowledge advancement (Denison 1962) and also as the "measure of our ignorance" (Abramovitz 1956). The large proportion of growth ascribed to this factor led to inquiries into its composition. See note 8.

4. A shortcoming of this process is that Denison attributes a proportion of economic growth to numerous sources that make sense theoretically but that may not represent direct inputs into the aggregate production function.
5. The term $\Delta X/X$ represents percentage changes, or growth rates in X.
6. See Denison (1962, 142, table 18). Denison (1974) refines the input measures, weighting the labor input, for example, to account for the distribution of education level and, in turn, for the increase in labor productivity. Despite the refinement, he still finds a large residual component—that is, that advances in exogenous technology contribute significantly to output growth.
7. Weisbrod (1961) estimates the capitalized value of human capital.
8. Note that Denison (1974) attributes 25 percent of U.S. growth since 1930 to the economic effects of education and still has more than 40 percent of U.S. output growth explained by the residual, which he interprets as the growth in knowledge.
9. Lucas (1988) and Stokey (1991) also suggest that external effects from human capital and knowledge accumulation may be a major source of economic growth.
10. Jorgenson and Fraumeni estimate the value of increased lifetime earnings related to further educational attainment by comparing the incomes of individuals that are the same age (and sex) but with different levels of education. This differential becomes the basis for a measure of the output of education. The procedure is similar to Denison (1962, 67-69).
11. Jorgenson and Fraumeni (1991) conclude that their labor input measure accounts for more than 61 percent of the estimated growth in U.S. output; the capital input accounts for more than 22 percent, and the residual (technology) input accounts for only about 17 percent. Theirs are time-series estimates versus Denison's long-run averages.
12. Easterlin (1981) appears to take this perspective toward developing nations but also hints at external effects of education by discussing the growth in knowledge separate from educational investment.
13. The analysis and graph are taken directly from Sala-i-Martin (1990a, 1990b). The rate of return equation is often referred to as the Keynes-Ramsey equation.
14. Useful surveys of the endogenous growth literature include, for theory, Sala-i-Martin (1990b) and van de Klundert and Smulders (1991). See Renelt (1991) for a review of the empirical work as well.
15. Romer's empirical work, which will be examined later, investigates his contention that the initial level of human capital may explain subsequent growth (similar to the ideas of Easterlin 1981).
16. This discussion, of course, fails to convey the technical difficulty of creating these models and deriving the results. For the purposes of this study, however, the main results and implications are the relevant portions of the papers.
17. Romer (1990) suggests that this finding may be the result of measurement error. Azariadis and Drazen (1990) do not include the rate of investment in their regressions to examine whether their human capital proxy would retain its significance in the output regression.
18. They use the average (from 1960 to 1985) percentage of secondary school-aged individuals in school as the human capital proxy. The output variable is the natural logarithm of output per working-aged individual.
19. They employ a version of Leamer's variance bounds tests that examines how much a coefficient estimate changes when the set of explanatory variables is altered (see Leamer 1978). The variables that they examine include the 1960 literacy rate used in Romer (1990) and Azariadis and Drazen (1990) as well as the 1960 levels of primary and secondary school enrollment rates used in Barro (1991).
20. Benhabib and Speigel (1991) provide a recent exception.
21. The study examines several different weighting schemes, in each assigning greater weight to better-educated workers (assuming more productive workers). Additionally, a consistent capital stock time series is created using the estimated capital stock from 1975 and the aggregate investment series.
22. Lucas (1988) suggests that the growth rates of countries may converge but there is no sense of convergence in the level of output per capita.
23. Levine and Renelt (1992) find a convergence result that appears robust in a specification that includes a human capital measure in the regression.
24. Quah (1990) emphasizes the problem known as Galton's fallacy of regression toward the mean. The tests use the output level at the beginning of a period as an explanatory variable for the subsequent output growth rate. Quah shows that the estimated coefficient can be negative, positive, or zero for the same cross-sectional distribution. As a result, the sign of the estimated coefficient of initial levels provides no information about whether the cross-section of country outputs converges or diverges. The negative coefficient on initial output level found in the studies cited above is viewed as evidence in support of convergence.
25. Learning on the job and specific job training, additional important sources of human capital investment, appear more difficult to examine empirically.
26. Although some research (cited above) argues for a better allocation of educational effort toward primary education, there is no call for a net reduction in total education level.

References

Abramovitz, Moses. "Resources and Output Trends in the United States since 1870." *American Economic Review Papers and Proceedings* 46 (May 1956): 5-23.

Azariadis, Costas, and Allen Drazen. "Threshold Externalities in Economic Development." *Quarterly Journal of Economics* 105 (May 1990): 501-26.

Barro, Robert J. "Economic Growth in a Cross-Section of Countries." *Quarterly Journal of Economics* 106 (May 1991): S103-25.

_____, and Xavier Sala-i-Martin. "Convergence." *Journal of Political Economy* 100 (April 1992): 223-51.

Becker, Gary. *Human Capital*. Chicago: University of Chicago Press, 1975.

Benhabib, Jess, and Mark M. Speigel. "Growth Accounting with Physical and Human Capital Accumulation." New York University, photocopy, November 1991.

Denison, Edward F. *The Sources of Economic Growth in the United States and the Alternatives before Us*. New York: Committee for Economic Development, 1962.

_____. *Accounting for United States Economic Growth, 1929-1969*. Washington, D.C.: The Brookings Institution, 1974.

Easterlin, Richard A. "Why Isn't the Whole World Developed?" *Journal of Economic History* 61 (March 1981): 1-19.

Jorgenson, Dale W., and Barbara M. Fraumeni. "Investment in Education and United States Economic Growth." Harvard Institute of Economic Research Discussion Paper 1573, October 1991.

Kendrick, John W. *Postwar Productivity Trends in the United States, 1948-1969*. New York: National Bureau of Economic Research, 1973.

Leamer, Edward E. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York: John Wiley and Sons, 1978.

Levine, Ross, and David Renelt. "A Sensitivity Analysis of Cross-Country Growth Regressions." *American Economic Review* 82 (September 1992): 942-63.

Lucas, Robert E. "On the Mechanics of Economic Development." *Journal of Monetary Economics* 22 (July 1988): 3-42.

Mankiw, N. Gregory, David Romer, and David N. Weil. "A Contribution to the Empirics of Economic Growth." *Quarterly Journal of Economics* 107 (May 1992): 407-37.

Nussbaum, Bruce, Aaron Bernstein, Elizabeth Ehrlich, Susan B. Garland, and Karen Pennar. "Needed: Human Capital." *Business Week*, September 19, 1988, 100-41.

Prescott, Edward C. "Theory Ahead of Business Cycle Measurement." Federal Reserve Bank of Minneapolis *Quarterly Review* (Fall 1986): 9-22.

Psacharopoulos, George. "The Contribution of Education to Economic Growth: International Comparisons." In *International Comparisons of Productivity and Causes of the Slowdown*, edited by John W. Kendrick, 335-55. Cambridge, Mass.: Ballinger Publishing Company, 1984.

_____. "Returns to Education: A Further International Update and Implications." *Journal of Human Resources* 20 (Fall 1985): 583-604.

Quah, Danny. "Galton's Fallacy and Tests of the Convergence Hypothesis." Massachusetts Institute of Technology, photocopy, 1990.

Renelt, David. "Economic Growth: A Review of the Literature." World Bank Working Paper WPS 678, May 1991.

Romer, Paul. "Human Capital and Growth: Theory and Evidence." *Carnegie-Rochester Series on Public Policy* 32 (1990): 251-86.

Rosen, Sherwin. "A Theory of Life Earnings." *Journal of Political Economy* 84 (August 1976): S45-67.

Sala-i-Martin, Xavier. "Lecture Notes on Economic Growth (I): Introduction to the Literature and Neoclassical Models." National Bureau of Economic Research Working Paper 3563, December 1990a.

_____. "Lecture Notes on Economic Growth (II): Five Prototype Models of Endogenous Growth." National Bureau of Economic Research Working Paper 3564, December 1990b.

Schultz, T. Paul. "Education Investments and Returns." In *Handbook of Development Economics*, 544-630. Amsterdam: North-Holland, 1988.

Schultz, Theodore W. "Investment in Human Capital." *American Economic Review* 61 (March 1961): 1-17.

Silk, Leonard. "Riots Put Focus on Economic Ills." *New York Times*, May 6, 1992.

Solow, Robert M. "A Contribution to the Theory of Economic Growth." *Quarterly Journal of Economics* 70 (February 1956): 65-94.

_____. "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics* (August 1957): 312-20.

Stokey, Nancy. "Human Capital, Product Quality, and Growth." *Quarterly Journal of Economics* 106 (May 1991): 587-616.

Swan, T.W. "Economic Growth and Capital Accumulation." *The Economic Record* 32 (November 1956): 334-43.

Tallman, Ellis W., and Ping Wang. "Human Capital and Endogenous Growth: Evidence from Taiwan." Federal Reserve Bank of Atlanta unpublished paper, 1992.

Van de Klundert, Theo, and Sjak Smulders. "Reconstructing Growth Theory: A Survey." CentER, Tilburg University, Working Paper 9146, July 1991.

Weisbrod, Burton. "The Valuation of Human Capital." *Journal of Political Economy* 69 (October 1961): 425-36.

Regional Employment By Industry: Do Returns to Capital Matter?

Stacy E. Kottman

States and communities in the United States have long pursued economic development. A variety of programs, including tax incentives, direct subsidies, and advertising, have met with mixed success over the years. Whatever the strategy—recruiting firms from other regions, promoting new foreign investment, or attempting to stimulate local entrepreneurship—the goal of these communities is to influence the regional distribution of industries across the country. Effective development efforts, those that improve a region's long-term relative wealth and well-being, should be grounded on an understanding of the economic forces that shape an area's evolving industrial composition.

Stories about regional growth and interregional shifts in industrial composition do not make the news as frequently as in the heyday of competition between the Rust Belt and Sun Belt during the 1970s. Nonetheless, regional shares of national employment continue to change, with the extent of the changes varying widely from industry to industry regardless of what is happening to total industry employment at the national level. These movements of employment and capital among regions are critical in explaining differences in regional growth rates. In addition, regional shifts in industrial composition play a key role in determining changing patterns of per capita income and labor productivity.

Several fundamental cause and effect questions surround the issue of interregional industrial migration: In an economy in which capital and labor have long been able to migrate across regions without restriction, why do interregional employment and capital shifts continue? Is this migration a form of arbitrage, exploiting fundamental economic imbalances among regions? Or are U.S. regions generally in economic equilibrium only to be

The author is an economist in the regional section of the Atlanta Fed's research department. *He is grateful to Ellis Tallman and Tom Cunningham for helpful comments and Hing Kwong Law for indefatigable research assistance.*

buffeted by economic "shocks" with region-specific effects, such as dramatic changes in oil prices, defense expenditures, or state development initiatives?

From the viewpoint of conventional economic models used to study regional growth, it is puzzling that interregional capital flows have persisted despite the fact that workers appear content to receive different wages as compensation for regional differences in cost of living or amenities (that is, in technical terms, labor markets are in equilibrium across regions).[1] The continued interregional employment shifts under these conditions suggest that relative profitability—differentials in returns to capital—may be the source of the capital flows.

If so, there are implications not only for theoretical models of regional growth and development policy but also for regional econometric models used to forecast economic activity or evaluate the impact of development initiatives. In conventional regional econometric models, changes in industry employment are primarily dependent on changes in national or local product demand. These top-down, demand-oriented models typically assume away the potential impact of regional differentials in returns to capital (see the box on page 17) and, consequently, disregard interregional capital flows, which may be more important to long-run regional growth and development than existing industrial composition.

The purpose of this study is to address the fundamental question of whether regional differentials in returns to capital drive changes in regional employment. Empirical tests assess the influence of relative profitability on changes in industry employment across U.S. states from 1969 through 1986.

The results confirm that for industries producing goods and services for export to other regions there is a positive, significant relationship between relative returns to capital and a state's share of national employment. Regional differentials in profitability were not found to be significant in explaining employment share for industries that produce goods or services for local consumption. For these industries local market size is the key determinant of employment share over time.

Interregional Mobility by Industry

In this study interregional mobility is broadly defined as the rate at which regional employment shares—industry ratios of regional employment to national employment—change over time.[2] (In any year the sum of regional shares for an industry will equal 1.0.) Considering regional employment mobility in these terms largely neutralizes the influence of business cycles and the longer-term effects of changes in technology or international competition. For instance, in a declining industry in which employment is falling faster nationally than in a region, the region will gain national employment share.

An example contrasting employment mobility in four industry categories in two U.S. regions will illustrate the concept. Using Bureau of Economic Analysis (BEA) regions, national employment shares for the Mideast and Southeast are calculated for total employment, services, retail trade, and electronic equipment over the 1969-90 period.[3] These regions were selected because they had similar total employment levels in 1969 and exhibited divergent employment trends. Employment share for each category is plotted in Chart 1.

Over the twenty-one-year period, total employment share drifted steadily lower in the Mideast and higher in the Southeast. The relative shift in services was smaller than for total employment in both regions. In fact, the Southeast's share of national service employment has remained relatively steady. In retail trade, the regional shifts proceeded in the same direction as total employment but at a modestly faster pace.

At the same time, a dramatic shift occurred in national employment share for the electronic equipment industry. In 1969 the Mideast employed 24 percent of the nation's electronic equipment workers, more than twice as many as the Southeast. By 1990, however, the Southeast's electronic equipment work force was one-third larger than the Mideast's. Mideast employment levels fell nearly in half while the Southeast's electronic equipment work force rose by 44 percent.

To assess whether such variations in industry mobility are observable across all eight regions, Table 1 offers a simple aggregate measure of interregional mobility for thirty-four private-sector industries from 1969 to 1990. By industry, national employment shares for the eight BEA regions were calculated for 1969 and 1990.[4] For each region, 1969 employment share was subtracted from 1990 employment share. The absolute value of these regional employment share differences were then summed by industry.

The resulting measure is a crude but revealing index of interregional mobility across industries. By design, the index does not provide information on industry growth, contraction, or cyclical performance. It does offer a comparable measure of net regional employment mobility for industries in national decline as well as expansion.

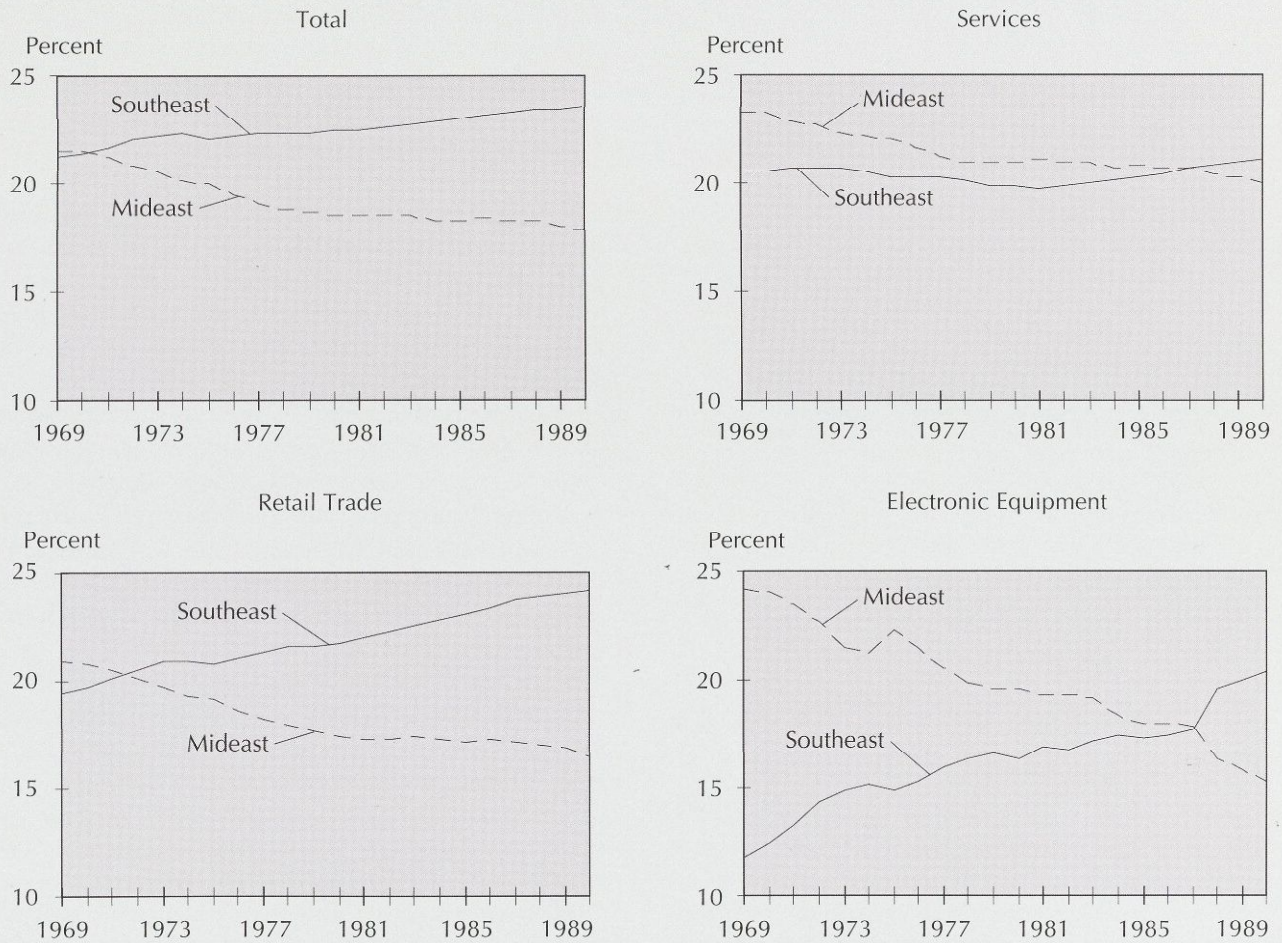
Two regularities are immediately apparent from Table 1. First, manufacturing industries are the most mobile, with their higher numbers dominating the bottom half of the ranking, while services, trade, and utilities are the least mobile, appearing in the top half of the list. Second, the manufacturing industries that exhibit relative immobility are typically restrained by what economists call fixed effects. For example, the production of chemicals, lumber, and food products all depend heavily upon geographically fixed raw resources.

To the extent that these employment shifts serve as a proxy for regional investment, this evidence suggests that interregional capital flows vary systematically by type of industry. Empirically identifying the fundamental economic incentives that determine this variable migration requires a conceptual framework, or economic model.

Interregional Capital Migration

Do Returns to Capital Matter? Conceptually, it is neither innovative nor imaginative to assume that capital will seek its highest return. It is relative profitability and changes in the regional profit topology that interest owners of capital (James A. Chalmers and Terrance L. Beckhelm 1976). Unfortunately, because of limited data on regional capital stock and returns to capital, the determinants and pattern of capital migration have

Chart 1
Employment Share in the Mideast and Southeast, 1969–90
(Total and by selected industries)



Source: Calculated by the Federal Reserve Bank of Atlanta using Bureau of Economic Analysis data.

Table 1
Regional Employment Mobility by Industry

Mobility Index[a]	Industry	BEA Industry Code
0.115	Chemicals and Allied Products	260
0.133	Lumber and Wood Products	320
0.134	Health Services	621
0.137	Personal Services	580
0.147	Wholesale Trade	510
0.156	Amusement Services	611
0.162	Construction	180
0.168	Insurance Agents and Brokers	554
0.177	Retail Trade	520
0.185	Electric, Gas, and Sanitary Services	500
0.187	Communication	490
0.193	Hotels and Lodging	570
0.196	Food and Kindred Products	210
0.196	Trans. Equip., except Motor Vehicles	380
0.197	Trucking and Warehousing	460
0.201	Legal Services	622
0.201	Paper and Allied Products	240
0.205	Furniture and Fixtures	330
0.225	Textile Mill Products	220
0.252	Business Services	601
0.258	Fabricated Metal Products	350
0.258	Transportation by Air	482
0.261	Real Estate Services	555
0.267	Stone, Clay, and Glass Products	410
0.276	Printing and Publishing	250
0.278	Motor Vehicles and Equipment	390
0.288	Primary Metals Industries	340
0.308	Transportation Services	484
0.337	Miscellaneous Manufacturing Industries	430
0.353	Machinery, except Electrical	360
0.375	Rubber and Miscellaneous Products	290
0.376	Electric and Electronic Equipment	370
0.432	Apparel and Other Textile Products	230
0.510	Instruments and Related Products	420

[a] *A low mobility index suggests relatively less interregional change in employment share across eight U.S. regions during the 1969-90 period. The index is reported for those industries that are evaluated empirically in the article.*

Source: Calculated by the Federal Reserve Bank of Atlanta using Bureau of Economic Analysis data.

received relatively little attention compared with the voluminous research on wages and interregional labor migration. Additionally, data constraints have limited research on interregional capital flows to the manufacturing sector.[5] However, existing empirical research provides an important foundation from which this study proceeds.[6]

Robert F. Engle (1974) regressed an approximation of relative returns to capital on an estimate of total investment in Massachusetts for four highly aggregated manufacturing sectors over a sixteen-year period. Although his results confirm a positive and significant role for differentials in returns to capital, he did not test his proposition across a cross-section of states.

Lynne E. Browne, Peter Mieszkowski, and Richard F. Syron (1980) concluded that the South's rapid capital expansion was not attributable to the existing industrial composition but rather to interregional capital migration. In addition, their study indicates that low nominal wages played the most important role in attracting net investment into the South. Using surveys, other researchers have confirmed that low wages and a low level of unionization are important factors in investment location decisions (Robert A. Nakosteen and Michael A. Zimmer 1987; Leonard F. Wheat 1986; Robert J. Newman 1983; W.C. Carlton 1979). These surveys also reveal that proximity to markets is a key factor in location decisions.

Edward M. Miller (1981) tested whether nominal wage differentials compensate for differences in labor productivity. If so, unit labor costs could be in equilibrium across regions despite nominal wage differentials. He found that variations in labor productivity only partially offset wage differentials. Not surprisingly, after adjusting for the differences in labor productivity, Miller found a positive correlation between lower wages and profitability. However, he makes no attempt to correlate capital migration to estimates of relative profitability.

A synthesis of these results suggests that capital migration does respond to profit differentials that may originate from regional variation in production costs. Assuming that product price and the cost of capital are nationally determined, imbalance in regional wages appears to play a primary role in determining relative regional profitability and, by extension, interregional capital flows in the manufacturing sector.

However, two issues remain unresolved. These results do not immediately explain why the pace of interregional capital migration would vary widely among manufacturing and nonmanufacturing industries. More intriguing is the question of how these results square with the growing body of evidence that the regional labor markets are in equilibrium.

How Firms and Labor View Regional Wages. There is an emerging consensus in the literature that regional nominal wage and income differentials have all but vanished once adjusted for cost-of-living and

Labor Demand in Regional Econometric Models

While excellent labor market data are available, the absence of consistent capital stock, net investment, and output data has rendered direct production function estimation for regional models nearly impossible. In 1972 Donald Ratajczak proposed a methodology to specify state-specific labor demand without capital stock data. This methodology has since become widely used. Assume that (1) the production relationship can be approximated by constant returns to scale;[1] (2) firms exhibit profit-maximizing behavior in perfect competition; (3) the labor market is in equilibrium; and (4) regional returns to capital are in equilibrium.

Under these assumptions, it can be shown that labor is paid its marginal product and the derived demand for labor can be estimated without capital data. The legitimacy of this theoretical specification rests heavily on its assumptions. Ratajczak (1972) notes that differential capital returns, or profitability, could stimulate interregional capital flows. However, the consideration of disequilibrium is dismissed because of data constraints.

In practice, the conventional labor demand is presented as follows:

$$l_n L_t = a + b(l_n Q_t) + c(l_n W_t) + d(T_t),$$

where l_n is log, L_t is demand for labor services in time t, Q_t is output (national or local demand or both), W_t is the real wage, and T_t is the level of technology (with $l_n L_{t-1}$ often added as a partial adjustment term) (Max E. Jerrell and James M. Morgan 1988; Roger Bolton 1985; William J. Milne, Norman J. Glickman, and F. Gerard Adams 1980; Ratajczak 1972).

The limitation of this equilibrium specification in determining relative regional employment growth among states can be seen by examining the paths through which relative labor demand could be expressed. Under equilibrium conditions, changes in aggregate industry demand would result in proportional, not relative, changes in labor demand.

Technology or innovation would not affect relative labor demand by industry unless it was region-specific. The assumption of limited diffusion is not credible in today's economy unless it is dependent on a region-specific resource, such as raw material. Generally, technology is assumed to displace labor. Therefore, the expected sign of d is negative.

The real wage, which in most regional models is simultaneously dependent on labor demand and supply conditions, inversely influences labor demand in the standard specification. Therefore, the coefficient c should be negative. In fact, empirical estimation of this specification often results in a positive wage coefficient (Gordon L. Clark, Meric S. Gertler, and John E. Whiteman 1986).

Adjusted for aggregate demand, change in the derived demand for labor under this specification is limited to wage-driven capital-labor substitution and technological change. The equilibrium specification does not consider employment shifts driven by differentials in interregional returns to capital.

Note

1. Technically, a constant elasticity of substitution (CES) production function, homogeneous of degree one, with disembodied technological change is specified for estimation purposes.

labor force characteristics (see John A. Bishop, John P. Formby, and Paul D. Thistle 1992; Gerald A. Carlino 1992; Mark Dickie and Shelby Gerking 1987; Gary M. Fournier and David W. Rasmussen 1986; Leonard G. Sahling and Sharon P. Smith 1983; Don Bellante 1979). The labor market, from the viewpoint of a worker, appears to be in equilibrium.

However, industries continue to migrate across regions in response to profit differentials generated by the apparent lack of wage equilibrium. The conundrum lies in the fact that real wages from labor's viewpoint are not equivalent to real wages from the viewpoint of the firm (see Daniel H. Garnick and Howard L. Friendenberg 1982; Stacy E. Kottman 1990; Carlino 1992). From the viewpoint of labor, for example, a 10 percent wage differential may compensate for regionally determined differences in cost of living or amenities. However, to a firm producing a product whose price is determined nationally, a 10 percent wage differential may represent a 10 percent differential in labor costs (assuming similar labor productivity across regions).

For the firm, labor cost differentials may create an opportunity for more profitable production in the low-wage region and may drive capital across regions. Whether this migration results in a narrowing of regional wage differences depends on the resultant change in relative industrial composition. In fact, Browne, Mieszkowski, and Syron (1980) concluded that the Northeast had become relatively more concentrated

in high-wage, or capital-intensive, industries as lower-wage, or labor-intensive, production moved South. Incorporating this insight helps to explain the remaining unresolved issue—the varying rate of interregional mobility across industries.

Expectations of Mobility Vary by Industry. Conceptualize the following industrial spectrum. Toward one end is an industry—the neighborhood dry cleaner, for example—whose total regional capital investment and employment is directly proportional to local market size. The product or service is produced primarily for local customers, and both input prices, such as labor costs, and output prices are determined by local markets. If it is assumed that capital requirements for firm entry are not prohibitive, competition quickly squeezes out opportunities to set high prices and raise profits. In such an industry, differentials in returns to capital, or profitability, would not vary appreciably across regions. Therefore, once adjusted for local market size, locally oriented industries do not engage in interregional migration in search of higher returns to capital.

Toward the opposite end of the spectrum is an industry like automobile assembly that, after using local inputs, exports all production. Product demand and price are largely determined in national markets. Although the cost of capital is determined nationally, other important input prices—primarily wages—are determined locally and could vary across regions, allowing regional differentials in total returns to capital. This relative profitability most likely would initiate interregional capital flows resulting in shifting regional employment shares. Local market size should by comparison be unimportant in determining a region's employment share for export-oriented industries.

All thirty-four industries in Table 1 lie somewhere along this conceptual spectrum. A casual evaluation of industry mobility as ranked by the table suggests that export-oriented manufacturing industries do exhibit more mobility than locally oriented industries such as services, retail, and utilities. To test this proposition on an industry-by-industry basis, an empirical model was estimated to examine the importance of relative returns to capital as a determinant of employment share across states from 1969 through 1986.[7]

Model Specification

The influence of relative returns to capital is modeled for each industry as follows:

$$ES_{jt} = a + b(RRC_{jt-1}) + c(MS_{jt-1}) + d(FXD_j) + e_{ij},$$

where j represents the state, t is the year, ES is national employment share, RRC is relative returns to capital, MS is share of national market, FXD specifies state fixed effects, and e is random error.

For any industry this model specifies that state employment share (ES_{ij}) is a lagged function of relative returns to capital (RRC_{ij}), state market size (MS_j), and state fixed effects (FXD_j) that are not lagged. This specification evaluates the role of relative profits for industries in absolute national decline as well as expansion. The model is not uniquely concerned about firms physically moving from state A to state B. By design, it considers relative regional shifts in net investment and employment, regardless of the source.

The dependent variable, ES_{ij}, was defined above, but a brief review of the explanatory variables, their empirical representations, and their expected sign or significance by industry should be helpful in interpreting the results.

Relative Returns to Capital. Conceptually, RRC_{ij} is defined as the ratio of state returns to capital to national returns to capital. The model assumes a lagged effect on employment share by considering the nature of capital investment. Once the decision for additional investment is made on the basis of observed changes in relative profitability, it takes time to locate a site, construct facilities, and hire a work force. Lags should vary with the complexity or fixity of the capital investment under consideration.

Unfortunately, no consistent annual corporate-profits or capital-stock series exists at the regional level. Previous empirical studies used value-added minus labor earnings as an admittedly imperfect proxy for total returns to capital in manufacturing industries. However, the BEA's recently released gross state product (GSP_{ij}) series by industry offers an opportunity to calculate a more serviceable measure of relative profitability across manufacturing and nonmanufacturing industries.

Gross state product by industry is defined as the sum of employee compensation, proprietors' income, indirect business taxes, and capital-related charges (BEA 1985, 1988, 1991; see the appendix for a detailed discussion). The BEA's estimate of capital-related charges (CRC_{ij}), which include adjusted corporate profits, rental income, net interest, and depreciation, is significantly more independent and consequently less distortive than the gross residual between value-added and labor earnings. Therefore, the ratio of capital-related charges to gross state product provides an

improved measure of relative returns to capital in each state. Because capital-related charges can be negative, relative returns to capital (RRC_{ij}) by industry was calculated as the difference between the state and national ratios in each year:

$$RRC_{ij} = (CRC_{ij}/GSP_{ij}) - (CRC_{ius}/GSP_{ius}).$$

If the model holds, the estimated coefficient for RRC_{ijt-1}, b, should be positive and significant for export-oriented industries. In fact, for export-oriented industries, differentials in returns to capital should be more important than local market share (MS_j). Relative returns to capital should not be significant for locally oriented industries, where local market share is expected to dominate.

Market Share. Surveys confirm that proximity to market is a key factor in location and investment decisions. Therefore, to reveal the true role of RRC_{ij}, the influence of local market share, MS_j, must first be accounted for. Local market share is calculated as the ratio of state to national total personal income in each year.

MS_j should be more significant for those industries that retain a local orientation in production, pricing, and consumption. Competitive pressures inherent in such industries would serve to limit persistent differentials in returns to capital among regions. The coefficient, c, for MS_j may be positive and significant for some export-oriented industries, but it should clearly dominate the results for locally oriented industries, where differentials in returns to capital are not expected to persist.

Fixed Effects. Fixed, but unobservable, effects—such as access to natural resources, climate, amenities, culture, political environment, and historical accident—are assumed to be significant in determining employment share by state. If such effects are assumed to exist and are correlated with other explanatory variables, ignoring their presence would introduce specification bias (Cheng Hsiao 1986). Because it is reasonable to assume that a state's fixed effects are correlated with relative returns to capital, they were accounted for by including dummy variables for each state.[8]

A major limitation of using dummy variables to capture fixed effects is that they do not identify which factors are responsible for the fixed effect. However, this characteristic has some value because it makes it possible to compare the significance of fixed-effect coefficients across industries and states, providing additional insight on mobility. State fixed effects are expected to be more important in some industries than others.

Estimation Issues

The Economic Importance of Relative Returns to Capital. Estimating the model with the relative ratios as defined may verify whether RRC_{ij} is statistically significant by industry but not whether it is economically important in comparison with local market share. Because of the wide cross-sectional variation in state sizes, employment share and market share will be highly correlated. Consequently, the estimated coefficient on MS_j should be statistically significant and large in value when the model is estimated with the ratios directly.

In contrast, the cross-sectional range of values for RRC_{ij} is relatively narrow. Although its statistical significance may be confirmed, the estimated coefficient on RRC_{ij} would likely be small relative to the coefficient on MS_j. It would thus be inappropriate to compare standardized coefficients for RRC_{ij} and MS_j in evaluating relative economic significance. However, estimating the model in first-difference form—that is, how the ratios change from year to year—and comparing the standardized coefficients provides an evaluation of the relative economic importance of each variable.[9]

State Sample Selection. GSP data are available for all fifty states and the District of Columbia. However, it is difficult to defend the proposition that these fifty-one areas represent independent samples drawn from a single population. In addition, several small states report a very small employment or GSP share. The smaller a state's share, the more vulnerable its data are to errors in measurement or dominant firms that raise disclosure problems (BEA 1985).

Because of these considerations, the model was estimated across a subset of the states.[10] The selection rule was simple: All states with less than a 1 percent share of national total manufacturing employment were deleted from the sample, leaving thirty states whose total manufacturing employment share summed to 93 percent of the national total. These larger states are more likely to share broad characteristics that make the underlying assumption of a single population more defensible, placing a reduced burden on the fixed-effects variables.

Industry Sample Selection. For estimation, industry selection was limited to private nonagricultural sectors. Two types of industries were excluded from testing on an a priori basis. First, industries with significant immobility were not considered: tobacco, petroleum and coal, mining, rail transportation, water

Table 2
The Significance of Relative Returns to Capital In Explaining Change in State Employment by Industry[a]

Industry	BEA Code	RRC_{ij}	MS_j	$RRC_{ij} > MS_j$
Export-Oriented				
Primary Metals Industries[b]	340	+HS	−I	Y
Fabricated Metal Products	350	+HS	+I	Y
Electric and Electronic Equipment	370	+HS	+I	Y
Motor Vehicles and Equipment[b]	390	+HS	−I	Y
Chemicals and Allied Products[b]	260	+HS	+S	Y
Miscellaneous Manufacturing Industries	430	+HS	+I	N
Instruments and Related Products	420	+S	+HS	N
Machinery, except Electrical	360	+S	−I	Y
Trans. Equip., except Motor Vehicles[b]	380	+S	−I	Y
Textile Mill Products[b]	220	+S	+I	Y
Lumber and Wood Products	320	+I	+I	N
Furniture and Fixtures	330	+I	−I	Y
Apparel and Other Textile Products	230	−I	−I	N
Rubber and Miscellaneous Products	290	−I	−I	N
Both Export- and Locally Oriented				
Stone, Clay, and Glass Products[b]	410	+HS	−I	Y
Transportation by Air	482	+I	+S	N
Printing and Publishing	250	+I	+S	N
Paper and Allied Products	240	−I	+HS	N
Food and Kindred Products	210	−HS	+I	N
Locally Oriented				
Legal Services	622	+HS	+HS	N
Construction	180	+S	+HS	N
Real Estate Services	555	+I	+HS	N
Health Services	621	+I	+HS	N
Insurance Agents and Brokers	554	+I	+HS	N
Hotels and Lodging	570	+I	+HS	N
Electric, Gas, and Sanitary Services	500	−I	+HS	N
Trucking and Warehousing	460	−I	+HS	N
Wholesale Trade	510	−I	+HS	N
Retail Trade	520	−I	+HS	N
Personal Services	580	−I	+HS	N
Communications Services	490	−HS	+HS	N
Amusement Services	611	−HS	+S	N
Transportation Services	484	+I	+I	N
Business Services	601	−I	−I	N

[a] *This table summarizes the first-difference estimation results for thirty states. It shows the sign and level of statistical significance for the lagged RRC_{ij} coefficient (HS = highly significant [significant at a 1 percent level]; S = significant [significant at a 5 percent level]; I = insignificant), the sign and level of significance for the lagged MS coefficient, and whether the standardized coefficient for RRC_{ij} is larger than MS_j.*

[b] *For these industries, changes in RRC_{ij} were not significant in explaining changes in employment share. Reported results reflect regressing RRC_{ij} in level form on changes in employment share.*

Source: Calculated by the Federal Reserve Bank of Atlanta using Bureau of Economic Analysis data.

transportation, and pipelines. In these industries fixed effects in the form of natural resources or capital fixity limit potential migration. Several other industries were excluded from consideration on the basis of data considerations.[11]

Aside from these deletions, data were prepared across thirty-four private, nonagricultural industries. Because many industries exhibited some degree of missing or disclosure data problems, a sample of states with complete data was identified by industry.[12] The specified variables were then calculated and arranged into a cross-section time-series format over the 1969-86 period for estimation.

Results

The results confirm that employment share responds in a dynamic fashion to differentials in lagged returns to capital for those industries that are export-oriented. Table 2 presents summarized results across all thirty-four industries, categorized as export-oriented, locally oriented, or industries that should exhibit characteristics of both.[13]

Fourteen manufacturing industries are labeled export-oriented. After adjustment for local market share and fixed effects, RRC_{ij} was found to be statistically significant in determining employment share in ten industries. In eight of these ten, the standardized coefficient on relative returns to capital was larger than the standardized coefficient on market size.

An examination of the state dummy variable coefficients for the four export-oriented industries for which RRC_{ij} was not significant confirms a dominant role for state fixed effects. Natural resource fixed effects no doubt play a direct major role in lumber and wood and an indirect role in furniture and fixtures and rubber and plastics. Apparel reflected fixed effects in the form of high preexisting employment concentrations over few states.

Five industries were thought to exhibit both export and import characteristics. RRC_{ij} was only significant in stone, clay, and glass while MS_{ij} was significant in air transportation services, printing and publishing, and paper and allied products. Neither RRC_{ij} nor MS_j were able to explain employment share in food and kindred products, which exhibited significant fixed effects.

Sixteen individual industries were considered locally oriented. In contrast to export-oriented industries, MS_j was significant in all but two cases. Even though RRC_{ij} was significant for two industries (legal services and construction), the standardized coefficient for market share was larger in every industry, substantiating the relative importance of market share for determining employment share in locally oriented industries. In general, these results confirm that regional differentials in returns to capital are not important in determining interregional employment shifts for locally oriented industries.

Conclusion

This empirical study examines a simple but fundamental question regarding the interregional migration of capital and employment. Do differentials in returns to capital matter? For each of thirty-four industries, a fixed-effects model was estimated with pooled data over eighteen years and thirty states to evaluate the role of differentials in profitability on relative employment growth. Appropriate proxies for returns to capital were constructed from gross state product data by industry, and an appropriate estimation procedure was applied.

To the extent that capital is not influenced by non-economic fixed effects, the results suggest that differentials in returns to capital exert a significant and positive dynamic influence on interregional employment growth for export-oriented industries. For those industries, differentials in returns to capital are generally more important than local market size in explaining changes in state employment share over time.

As expected, entry and competition preclude a lagged dynamic relationship between returns to capital and relative employment growth for locally oriented industries. In these cases, employment share is primarily a positive function of local market size.

The results are consistent with the proposed theoretical framework that firms and workers hold differing viewpoints on nominal regional wage differentials. From the viewpoint of labor, the regional labor markets appear to be in general equilibrium. Nominal wage differentials simply compensate for differences in regional cost-of-living or amenities. However, from the viewpoint of the export-oriented firm, lower nominal wages may offer a more profitable investment opportunity that motivates interregional migration of capital and employment.

These findings yield several implications for regional development policy and regional growth models. First, low-wage states should not count on long-term

convergence to national per capita norms of income and output. Although conventional regional growth models maintain that convergence should occur as a result of interregional migration, the conceptual framework and empirical evidence presented in this article suggest an alternative scenario. Employment in low-wage regions could grow faster than in the nation without exhibiting wage convergence by absorbing increasing shares of the nation's relatively labor-intensive, low-wage industries. These results confirm the importance of looking beyond aggregate per capita measures of incomes and output when discussing regional dynamics.

Second, interregional capital and employment mobility is a rational response to persistent differentials in returns to capital across regions for export-oriented firms. As evidenced by the Mideast/Southeast comparison, these interregional shifts in industrial composition have been relatively steady throughout the business cycles and oil shocks of the past twenty-one years. Broadly interpreted, this persistent interregional migration represents a search for regional equilibrium that may be more pervasive than adjustments to periodic exogenous price shocks.

One sobering extension of this interpretation for development policy is that marginal tax incentives or investment subsidies may have little effect on a region's employment composition over time. While development officials are quick to claim victory when recruiting firms, the long-term costs and benefits of direct subsidies are not so clear. In fact, direct investment incentives could conceivably deter state progress toward national norms by subsidizing inefficient firms or by displacing public investment in policies that could truly lift a region's relative per capita earnings.

For regional economists engaged in econometric modeling, these results suggest that specification bias may lurk in conventional specifications of labor demand by industry, which assume away differentials in returns to capital as a source of labor demand. Although this potential specification bias may present few problems in short-term forecasting, it could be troublesome when using demand-driven models for longer-term policy evaluation.[14]

Appendix
Gross State Product and Other Data Series

In 1988 the BEA reported gross state product (GSP) for each of sixty-one industries at the two-digit Standard Industrial Classification (SIC) level from 1963 through 1986. GSP is defined as the sum of four independently generated and reported component series (BEA 1985, 1988, 1991):

1. Compensation of employees, which includes wages, salaries, employer contributions for social insurance, and other labor income.
2. Proprietors' income with inventory valuation and capital consumption adjustment, which is the income, including income-in-kind, of sole proprietorships, partnerships, and tax-exempt cooperatives.
3. Indirect business taxes (IBT_{ij}) and nontax liabilities, which consist of tax liabilities that are chargeable to business expense in the calculation of profit-type income. Normally, this category would include sales, excise, and property taxes and regulatory or inspection fees. It does not include corporate income taxes.
4. Capital-related charges (CRC_{ij}), which consist of
   (a) corporate profits with inventory valuation adjustment, which is the income of corporations measured before profit taxes, before deduction of depletion charges, after exclusion of capital gains and losses, and net of dividends received from domestic corporations;
   (b) the rental income of persons from the rental of real property, imputed net rental income of owner-occupants of nonfarm dwellings, royalties received from patents, copyrights, and rights to natural resources;
   (c) net interest, which is the interest paid by business less interest received by it;
   (d) business transfer payments and subsidies less current surplus of government enterprises; and
   (e) capital consumption allowances, which are depreciation charges and the value of accidental damage to fixed business capital.

The BEA data differ from previous periodic Census Bureau estimates in four ways. First, the BEA adjusts the census product to include value added by central administrative offices. Second, the BEA subtracts the costs of purchased services from census estimates. Third, the BEA adjusts for differences in industrial classification between census and BEA sources for payroll data. Each of these adjustments relies on general distributional assumptions. However, any distortion caused by these adjustments should be smaller than the distortions that might exist in their absence. Fourth, the BEA includes nonmanufacturing industries.

In general, data for goods-producing and regulated industries offer a higher degree of independence and a lower reliance on distributional assumptions. More extensive assumptions are required for calculating service and trade industry components. Nonetheless, in the absence of complete information these series offer a defensible approximation of gross product originating by industry. The potential data distortions should not nullify the inference.

The BEA's new capital-related charges CRC_{ij} series is significantly more independent and less distortive than the gross residual between value-added and labor earnings, which functioned as a CRC_{ij} in previous empirical studies. The CRC_{ij} series excludes IBT_{ij}, which by definition was included in earlier residual estimates. To the extent that IBT_{ij} varies across states, its exclusion further reduces previous distortions.

In this article, capital-related charges (CRC_{ij}) are used as a proxy for total returns to capital. Does the inclusion of factors other than corporate profits inappropriately bias CRC_{ij} as a proxy for capital returns? Used in a comparative fashion across industries, the CRC_{ij} proxy could present potential distortions because the distributional assumptions that are used vary by industry. However, the hypotheses are to be tested on an industry-by-industry basis. To the extent that capital-labor ratios are similar within an industry across regions, CRC_{ij} is an unbiased proxy.

Market share (MS_{jt}) originates with the BEA's SA-5 personal income series (PI_{jt}) and is calculated as follows:

$$MS_{jt} = (PI_{jt}/PI_{ust}).$$

The BEA SA-25 series provided annual industry employment data. For estimation purposes, state employment share is calculated as follows:

$$ES_{ijt} = (EM_{ijt}/EM_{iust}).$$

Notes

1. Although economists have recently offered more sophisticated applications of neoclassical growth theory (see Barro and Sala-i-Martin 1992), these extensions remain variations on a simple neoclassical production model first applied to U.S. regions by Borts and Stein in 1964. Under standard neoclassical assumptions, firms in low-wage regions should exhibit a lower capital-to-labor ratio for any given production process, yielding a higher marginal value product for capital than is found in high-wage states. As capital migrates from the high-wage region to the low-wage region, capital-labor ratios in the low-wage region begin to rise, reducing returns to capital and increasing the marginal value product of labor. Consequently, wages and returns to capital will converge across regions. Once in equilibrium, interstate differences in employment growth rates for a given manufacturing industry arise solely from the wage impacts of relative shifts in regional labor supply.

   Borts and Stein concluded that low wages persist in some regions because of the intraregional migration of agricultural workers into manufacturing and the low-wage region's higher fertility rates. Although this explanation was consistent with demographic and industrial developments during the period under study (1919-57), it has become increasingly irrelevant since 1960 as the intraregional shift from agriculture to manufacturing has abated.
2. Changes in shares, the focus of this research, should not be confused with changes in actual employment levels.
3. The Mideast is defined as Delaware, the District of Columbia, Maryland, New Jersey, New York, and Pennsylvania. The Southeast encompasses Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, Virginia, and West Virginia. Employment data were taken from the BEA's SA-25 annual employment by industry series, which begins in 1969.
4. The BEA regions are Great Lakes, Mideast, New England, Plains, Southeast, Southwest, Rocky Mountain, and Far West.
5. State estimates of value-added, gross product, investment, or capital stock by industry typically have been restricted to the manufacturing sector.
6. The studies discussed are not intended to form an exclusive list. They were selected on the basis of their contribution to the questions posed by this article.
7. The author is currently reestimating the model with revised GSP data released in late 1991. The newly released data begin in 1977 and extend through 1989. Attempts are under way to link the new series with the original data, which extend back through 1963.
8. A random-effects model could introduce specification bias by ignoring the presence of fixed effects, so a covariance model is estimated that allows for differential intercepts over states (Hsiao 1986; Judge et al. 1985; Jakubson 1988). The FXD_{ij} dummy variable holds the value 1 for j and 0 for all other states to account for the effects of omitted variables that are specific to individual cross-sectional units but stay constant over time (Hsiao 1986). A joint F-test on the restricted versus unrestricted model was performed for each industry and confirmed that dummy variables provided important additional information. Because all variables are expressed as ratios of state-to-national values, time-specific effects, which would include national business fluctuations, interest rates, and inflation, have been discounted by variable

construction and need not be considered (Quan and Beck 1987). A joint *F*-test confirmed this assumption.

9. Because of data limitations, most empirical studies on this topic have been cross-sectional or time-series in nature and have not captured the dynamics of cross-sectional behavior over time. Although this model exploits new data and offers information on dynamics over time, it presents several intriguing econometric problems. In pooled estimation, disturbance terms could exhibit time-series correlation, heteroskedasticity, or cross-sectional correlation. Diagnostic tests from ordinary least squares estimation suggested first-order serial correlation and heteroskedasticity across all industries. Although serial correlation and heteroskedasticity will not bias estimators, it will bias standard errors and could lead to incorrect inferences from the application of significance tests. Therefore, a generalized least squares estimation was specified to correct for these error term problems, as in Kmenta (1971) and Johnston (1984). Usual panel data problems, such as sample selectivity, population inference, and limited degrees of freedom, are not issues with these data series.
10. Although this paper discusses results from the thirty-state sample, the model was also estimated across all forty-nine contiguous areas with consistent but slightly less robust results.
11. Industries were excluded on the basis of size and lack of geographic distribution (for example, leather manufacturing and motion pictures), assumed level of imputation on returns to capital (such as private household services and miscellaneous professional services), missing data, and where returns to capital can vary dramatically on the basis of changes in interest rates or financial market conditions (for example, holding and investment companies, banks and other credit agencies, and insurance carriers).
12. The BEA does not disclose state data when a firm's dominance in that state's industry makes its identification possible.
13. These categorizations are based on a priori expectations, not empirical evidence. Generally, manufacturing industries are considered export-oriented, and nonmanufacturing, locally oriented.
14. See Krikelas (1992) for a critique of economic base models.

References

Barro, Robert J., and Xavier Sala-i-Martin. "Convergence." *Journal of Political Economy* 100 (April 1992): 223-51.

Bellante, Don. "The North-South Differential and the Migration of Heterogeneous Labor." *American Economic Review* 69 (March 1979): 166-75.

Bishop, John A., John P. Formby, and Paul D. Thistle. "Convergence of the South and Non-South Income Distributions, 1969-79." *American Economic Review* 82 (March 1992): 262-72.

Bolton, Roger. "Regional Econometric Models." *Journal of Regional Science* 25 (November 1985): 495-520.

Borts, George H., and Jerome L. Stein. *Economic Growth in a Free Market*. New York and London: Columbia University Press, 1964.

Browne, Lynne E., Peter Mieszkowski, and Richard F. Syron. "Regional Investment Patterns." Federal Reserve Bank of Boston, *New England Economic Review* (July/August 1980): 5-23.

Bureau of Economic Analysis. "Experimental Estimates of Gross State Product by Industry." Staff Paper 42, 1985.

_____. "Gross State Product by Industry, 1963-86." *Survey of Current Business* 68, no. 5 (1988): 30-46.

_____. "Gross State Product by Industry, 1977-89." *Survey of Current Business* 71, no. 12 (1991): 43-59.

Carlino, Gerald A. "Are Regional Per Capita Earnings Diverging?" Federal Reserve Bank of Philadelphia *Economic Review* (March/April 1992): 3-12.

Carlton, W.C. "Why Do New Firms Locate Where They Do? An Econometric Model." In *Interregional Movements and Regional Growth*, edited by William C. Wheaton, 13-50. Washington, D.C.: The Urban Institute, 1979.

Chalmers, James A., and Terrance L. Beckhelm. "Shift and Share and the Theory of Industrial Location." *Regional Studies* 10, no. 1 (1976): 15-23.

Clark, Gordon L., Meric S. Gertler, and John E. Whiteman. *Regional Dynamics—Studies in Adjustment Theory*. Boston: Allen and Unwin, 1986.

Dickie, Mark, and Shelby Gerking. "Interregional Wage Differentials: An Equilibrium Perspective." *Journal of Regional Science* 27 (November 1987): 571-85.

Engle, Robert F. "A Disequilibrium Model of Regional Investment." *Journal of Regional Science* 14 (December 1974): 367-76.

Fournier, Gary M., and David W. Rasmussen. "Real Economic Development in the South: The Implications of Regional Cost of Living Differences." *Review of Regional Studies* 16 (Winter 1986): 6-13.

Garnick, Daniel H., and Howard L. Friendenberg. "Accounting for Regional Differences in Per Capita Personal Income Growth, 1929-79." *Survey of Current Business* 62 (September 1982): 24-34.

Hsiao, Cheng. "Analysis of Panel Data." Econometric Society Monograph, no. 11. Cambridge University Press, 1986.

Jakubson, George. "The Sensitivity of Labor-Supply Parameter Estimates to Unobserved Individual Effects: Fixed- and Random-Effects Estimates in a Nonlinear Model Using Panel Data." *Journal of Labor Economics* 6 (July 1988): 302-29.

Jerrell, Max E., and James M. Morgan. "Modeling Labor Demand in a State Econometric Model." *Review of Regional Studies* 18 (Fall 1988): 31-40.

Johnston, J. *Econometric Methods*. New York: McGraw-Hill, 1984.

Judge, George G., William E. Griffith, R. Carter Hill, Helmut Lutkepohl, and Tsoung-Chao Lee. *The Theory and Practice of Econometrics*. New York: John Wiley and Sons, 1985.

Kmenta, Jan. *Elements of Econometrics*. New York: Macmillan Publishing Company, Inc., 1971.

Kottman, Stacy E. "Relative Regional Employment Growth: The Role of Differentials in Returns to Capital." Ph.D. dissertation, Georgia State University, 1990.

Krikelas, Andrew C. "Why Regions Grow: A Review of Research on the Economic Base Model." Federal Reserve Bank of Atlanta *Economic Review* 77 (July/August 1992): 16-29.

Miller, Edward M. "Variation of Productivity, Wages, and Profitability with Location." *Review of Regional Studies* 11 (Spring 1981): 27-41.

Milne, William J., Norman J. Glickman, and F. Gerard Adams. "A Framework for Analyzing Regional Growth and Decline: A Multiregion Econometric Model of the United States." *Journal of Regional Science* 20 (May 1980): 173-89.

Nakosteen, Robert A., and Michael A. Zimmer. "Determinants of Regional Migration by Manufacturing Firms." *Economic Inquiry* 25 (April 1987): 351-62.

Newman, Robert J. "Industry Migration and Growth in the South." *Review of Economics and Statistics* 65 (February 1983): 76-86.

Quan, Nguyen T., and John H. Beck. "Public Education Expenditures and State Economic Growth: Northeast and Sunbelt Regions." *Southern Economic Journal* 54 (October 1987): 361-76.

Ratajczak, Donald. "An Econometric Specification of a Quarterly State Model." Unpublished manuscript, University of California–Los Angeles, 1972.

Sahling, Leonard G., and Sharon P. Smith. "Regional Wage Differentials: Has the South Risen Again?" *Review of Economics and Statistics* 65 (February 1983): 131-35.

Wheat, Leonard F. "The Determinants of 1963-77 Regional Manufacturing Growth: Why the South and West Grow." *Journal of Regional Science* 26 (November 1986): 635-59.

FYI

Tracking Manufacturing: The *Survey of Southeastern Manufacturing Conditions*

R. Mark Rogers

The Federal Reserve Bank of Atlanta, like the eleven other Reserve Banks across the nation, monitors economic conditions in its region. Probably its most important reason for doing so is to contribute to the Federal Reserve System's task of setting appropriate monetary policy. In the Southeast one of the most important influences on the economy's performance is manufacturing activity, which is more variable than most other sectors and accounts for a larger employment share in the region than it does in the nation. Consequently, to augment its analysis of current conditions in the region's economy, the Atlanta Fed's research department in the fall of 1991 launched the first comprehensive survey to focus solely on changes in indicators of manufacturing activity in the Southeast.

Because the Atlanta Fed's *Survey of Southeastern Manufacturing Conditions* has a rapid turnaround—less than three weeks for gathering, compiling, and reporting the data—it provides very recent information on the southeastern economy not available from other sources. During the past year the survey has proved increasingly valuable as it has been refined, the number of participants has steadily increased, and the patterns of responses have been better understood. In November the Atlanta Fed plans to begin releasing the survey data as a regular economic report.

This article traces the development and construction of the survey and explains the methodology used in compiling and calculating indexes from the data. The article also describes some uses of the survey data.

The author is forecast coordinator on the macropolicy team of the Atlanta Fed's research department. *He would like to thank Diana Cunningham, Cynthia Bansak, Teresa Beckham, James Mancuso, and Marguerite Lewis for their valuable contributions in the development of the survey.*

Developing the Survey

Before the southeastern manufacturing survey was developed, the Atlanta Fed monitored regional economic conditions chiefly by collecting anecdotal information and analyzing statistical data collected from various public sources. Approximately every six weeks the research department's regional team conducted an informal survey of business contacts to prepare a summary of current information about economic conditions in the Sixth Federal Reserve District. (This summary is still compiled with those from the other Reserve Banks in a publicly available document, *Current Economic Conditions by Federal Reserve District*, known as the Beigebook. The Atlanta Fed's summary is also used in briefing documents prepared for the Atlanta Bank president's use at Federal Open Market Committee [FOMC] meetings.)[1] Such contacts continue to provide valuable, up-to-the-minute information about regional conditions in all major sectors of the economy—not just manufacturing. Adding information from a formal manufacturing survey was seen as a way to corroborate and augment anecdotal information and provide a basis for comparing over time the producers' reports about various flows within the manufacturing process, such as orders, production, shipments, and inventories.

The survey was limited to the manufacturing sector for several reasons. Manufacturing activity is more readily quantifiable than that of many services industries. Moreover, because the manufacturing sector is one of the more cyclical components of the economy and is a key factor in pulling the overall economy into recovery or recession, timely information about conditions in the sector is critical for regional economic analysis. Finally, manufacturing is also important as an export base and as a high-wage sector, especially in the Southeast.

The methodology of the *Survey of Southeastern Manufacturing Conditions* (discussed in detail in the following sections) has a long tradition. Both the Philadelphia and Richmond Reserve Banks have produced manufacturing surveys for a number of years (see John Bell and Theodore Crone 1986; Christine Chmura 1987/88), and each is similar to the longer-running survey conducted by the National Association of Purchasing Management (1990). While the Atlanta Fed's survey generally adopted the methodology developed by the Philadelphia and Richmond Feds, it differs in several aspects. After refinements based on responses of a small test pool of plant representatives surveyed beginning in August 1991, the format was finalized in December 1991.

Sampling Criteria. All surveys are developed using some method of drawing respondents from the complete population. The Atlanta Fed's manufacturing survey used the manufacturing facility's location as the first criterion, choosing panelists at production facilities from across the six states that, in whole or in part, make up the Sixth Federal Reserve District.[2] Confining the survey to production facilities located in the Southeast eliminated companies headquartered in the region whose plants are elsewhere and, conversely, included plants operating in the region but owned by businesses headquartered in other parts of the country. This restriction makes the survey an indicator of activity in the Southeast exclusively.

Respondents also were chosen to reflect major industries in the broad range of manufacturing facilities within the region. They represent eighteen of the twenty industries classified as manufacturing industries according to the two-digit code of the Standard Industrial Classification (SIC) system (see Table 1).[3]

Dividing the Sample. In addition to classifying industries according to the two-digit SIC code, the Atlanta Fed survey's distribution of production units attempts to mimic the industry value-added distribution recorded in the 1987 Census of Manufactures.[4] ("Value-added" refers to the value added during the manufacturing process to the raw materials and intermediate products used as inputs.)[5] Currently, because its sample is small in relation to the number of classification categories, the survey distribution does not match exactly the Census of Manufactures' stratification. The Atlanta Fed survey weights each plant equally, and the distribution of respondents by industry varies somewhat from month to month. Table 1 shows the percentage distribution of respondents by industry for May-July 1992. As of September 1992, the survey was sent to approximately 250 manufacturing plants; the average response rate is slightly below 50 percent for the initial report.

Survey Information

The Atlanta Fed survey collects data on various types of manufacturing activity as indicators of trends, current activity, expectations, and plans. Manufacturers are first asked one question about the general performance of their firms' industry; the remaining questions relate to current and expected activity at their specific plants. Respondents are instructed to take seasonal variation into account and are assured of the confidentiality of their responses. They are questioned about

Table 1
State's Value-Added Manufactures as a Percentage of State's Total Value Added, 1987

SIC Code Number	Description	AL	FL	GA	LA	MS	TN	Region	U.S.	Survey Distribution by Units[a]
20	Food and kindred products	6.1	15.3	10.6	8.4	10.7	12.2	11.0	10.4	5.9
21	Tobacco products	D[b]	D	D	—	—	1.2	0.2	1.2	—
22	Textile mill products	7.8	D	D	D	D	3.2	1.7	2.2	4.9
23	Apparel and other textile products	7.1	3.1	5.6	1.2	7.2	5.6	4.9	2.8	4.7
24	Lumber and wood products	5.1	3.0	3.3	3.0	8.6	2.2	3.7	2.5	2.7
25	Furniture and fixtures	2.0	1.8	1.6	0.2	6.8	3.0	2.2	1.7	3.3
26	Paper and allied products	14.0	4.1	8.5	9.2	7.7	5.8	7.8	4.3	12.1
27	Printing and publishing	3.6	10.3	4.9	2.9	D	5.7	5.4	7.7	3.8
28	Chemicals and allied products	9.1	8.0	7.3	41.2	6.6	13.7	13.1	10.4	9.6
29	Petroleum and coal products	D	0.4	D	10.3	D	0.3	1.4	1.6	1.1
30	Rubber and misc. plastic products	6.4	2.5	2.4	0.9	4.5	5.8	3.7	3.8	4.9
31	Leather and leather products	—	D	D	—	D	1.0	0.2	0.4	—
32	Stone, clay, and glass products	2.8	4.3	3.3	1.6	2.8	3.5	3.2	2.9	2.5
33	Primary metal industries	8.6	1.2	2.9	0.9	2.7	3.1	3.1	4.0	2.7
34	Fabricated metal products	5.4	5.6	3.2	2.7	5.2	6.1	4.7	6.4	8.8
35	Machinery, except electrical	7.0	8.6	4.1	2.2	6.4	8.0	6.2	10.1	10.4
36	Electric and electronic equipment	5.6	12.6	7.1	4.9	8.9	6.3	7.7	8.2	11.2
37	Transportation equipment	6.0	7.0	14.1	9.2	9.3	8.3	9.4	11.8	9.9
38	Instruments and related products	1.3	10.2	1.5	0.2	0.6	2.0	3.1	6.1	1.6
39	Miscellaneous manufacturing	0.8	1.2	1.0	0.3	D	2.0	1.1	1.5	—
	Total	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

[a] *Average for May-July 1992.*
[b] D *indicates that data are withheld by Census because of disclosure regulation. Omission indicates that industry is not represented in that state.*

Source: U.S. Department of Commerce, Census of Manufactures, 1987.

production, shipments, new orders, order backlogs, materials inventories, finished goods inventories, number of employees, the average workweek, finished goods prices, raw materials prices, capital expenditures, new export orders, and supplier delivery time. The survey questions refer to three time frames: current month versus previous month, current month versus twelve months ago, and expected activity six months ahead compared with current levels. (See Table 2 for items and time periods.)

To minimize the reporting burden on panelists and to keep processing time to a minimum, the survey asks for the direction of changes for indicators rather than for specific levels of activity. For instance, instead of asking for the dollar volume of production for the current month, the survey asks whether production has decreased, increased, or remained unchanged.

Compiling the Data

The survey questionnaire requesting data pertaining to a particular month, along with the previous month's summary data, is mailed to manufacturers on approximately the twenty-fifth of the survey month. This procedure allows responses to be based on about three weeks of information concerning operations and on general knowledge about plans and expectations for the remainder of the month. Replies are requested to be returned by about the fifth of the following month. Each month, once the data are accumulated, the percentages of respondents showing improvement, deterioration, or no change for each activity indicator and time period are calculated. Final tallies for the initial survey release are completed by the tenth. Any late responses are incorporated into revised tallies. Summary statistics are revised only once and only to incorporate late survey responses. Release data are based solely on respondent information; there are no imputations or estimates for firms not replying.

Using the Data

The aggregate pattern of survey responses showing decrease, no change, and increase (see Table 2) provides information about changes in basic demand and supply flows as well as other data on items such as inflation and investments. This information should provide insight into manufacturers' current and expected performance. For example, data on monthly directions of change in new and unfilled orders, along with information on inventories, may indicate whether changes in production are likely. Data on directions of change in finished goods prices, raw materials prices, and supplier delivery time may serve as a rough inflation barometer. Information on directions of change in number of employees and the average workweek can provide insight into income trends as well as what plant managers believe about the degree to which production trends are sustainable.

Each of the survey's three time perspectives provides insight. Monthly comparisons, for instance, show progress over the current phase of the business cycle. Year-ago comparisons delineate longer-term trends and help indicate possible remaining seasonality in responses. Responses concerning expectations six months ahead (compared with current levels) can help ascertain whether recent changes in activity are expected to be temporary or not.

The regional economists and analysts of the Atlanta Fed's research department use the survey data to complement other information on the manufacturing sector. As the survey data suggest production trends, for example, other indicators, including regional employment and income as well as anecdotal evidence from contacts, can be monitored for confirmation.

Diffusion Indexes

To simplify determining how widespread changes are among respondents and comparing responses over time, the Atlanta Fed also calculates a statistical measure known as a diffusion index for each response category. A diffusion index is a tool used to gauge the similarity of individual reporter's changes in a particular category and time period.

The manufacturing survey's diffusion indexes—one for each time period of each indicator surveyed, such as production or new orders—are calculated as the percentage of total respondents reporting increases minus the percentage reporting declines.[6] (The percentage of "no-change" responses does not directly enter into the calculation although the number of "no-change" answers affects the other two percentages.) If all plants report increased activity for a particular category then the diffusion index for that category equals 100, and if all plants report declines then the category's value is −100. Thus, each diffusion index can range in value from 100 to −100. The logic of this method for calculating the diffusion index is that positive values

Table 2
Summary of Southeastern Manufacturing Conditions, July 1992
(Percentage of total respondents)[a]

	Current versus Previous Month			Current Month versus Year-Ago			Current Month versus Six Months Ahead		
	Decrease	No Change	Increase	Decrease	No Change	Increase	Decrease	No Change	Increase
Industry Business Conditions									
What is your evaluation of business activity in your industry?	25.0	51.5	23.5	25.0	23.5	51.5	15.7	27.6	56.7
Business Indicators									
Production	28.5	45.3	26.3	21.5	22.2	56.3	13.2	39.0	47.8
Volume of shipments	32.1	39.4	28.5	20.1	26.1	53.7	13.4	35.8	50.7
Volume of new orders	27.9	43.4	28.7	25.6	27.8	46.6	14.3	31.6	54.1
Backlog of orders	25.4	53.0	21.6	23.1	41.8	35.1	18.5	45.9	35.6
Inventories									
Materials	34.3	49.3	16.4	36.4	34.1	29.5	32.3	51.1	16.5
Plants' finished goods	35.9	45.0	19.1	39.5	28.7	31.8	32.3	46.9	20.8
Number of employees	18.8	60.9	20.3	34.1	29.6	36.3	18.2	53.3	28.5
Average employee workweek	16.1	67.9	16.1	17.9	50.0	32.1	13.2	69.1	17.6
Prices received for finished products	16.1	74.5	9.5	30.1	41.4	28.6	12.9	59.8	27.3
Prices paid for raw materials	8.3	73.5	18.2	13.4	35.8	50.7	4.4	54.8	40.7
Capital expenditures[b]				28.5	43.1	28.5	14.5	51.1	34.4
New orders for exports	9.5	75.9	14.7	10.8	64.2	25.0	2.5	70.5	27.0
Supplier delivery time[c]	6.9	87.0	6.1	10.9	71.9	17.2	3.9	79.8	16.3

[a] *Normal seasonal fluctuations are taken into account. Figures may not sum exactly to 100 because of rounding.*
[b] *Because firms plan capital expenditures on a long-term basis, this question is not applicable to month-ago comparisons.*
[c] *Decrease = slower; increase = faster.*

indicate increased activity at a majority of the plants over a certain time period while negative values indicate decreased activity at a majority of plants.[7]

Interpreting the Indexes. Although the Atlanta Fed survey does not collect data on specific levels of activity and cannot be used to deduce such levels, the diffusion indexes provide an indicator of the direction of changes in various types of manufacturing activity. Moreover, statistical studies indicate that over an extended period the level of diffusion indexes correlates well with growth rates in the economy (Ethan Harris 1991). See Table 3 for diffusion indexes

Table 3
Diffusion Index Summary of
Southeastern Manufacturing Conditions, July 1992

	1992							1991
	July	June	May	April	March	Feb.	Jan.	Dec.
Current versus Previous Month								
Industry Business Conditions	−1.5	37.4	18.2	35.4	34.8	14.6	3.8	−22.6
Plant Indicators								
Production	−2.2	33.6	16.8	29.0	40.4	24.1	21.8	−9.3
Volume of shipments	−3.6	33.0	13.4	27.3	39.4	19.5	11.5	−13.0
Volume of new orders	0.7	24.6	23.9	26.5	42.6	13.3	10.3	−20.4
Backlog of orders	−3.7	10.4	1.8	11.3	19.1	4.9	−7.9	−32.1
Inventories								
Materials	−17.9	−7.2	−3.6	−4.2	−3.3	3.8	−4.0	−11.5
Plants' finished goods	−16.8	1.8	0.9	−3.1	12.5	14.3	6.8	−13.7
Number of employees	1.4	15.4	1.8	13.3	5.4	3.7	−11.5	−11.3
Average employee workweek	0.0	18.8	2.7	16.0	8.6	10.0	1.3	−5.8
Prices received for finished products	−6.6	−0.9	3.6	0.0	−1.1	−2.6	−18.7	−5.9
Prices paid for raw materials	9.8	19.1	16.4	15.3	15.1	7.6	10.5	9.6
New orders for exports	5.2	14.3	9.8	21.2	17.1	8.3	11.7	4.3
Supplier delivery time[a]	0.8	8.0	0.0	2.1	−4.3	10.3	22.2	3.8
Current Month versus Six Months Ahead								
Industry Business Conditions	41.0	41.5	49.1	54.2	59.6	60.0	45.5	47.2
Plant Indicators								
Production	34.6	30.9	40.2	53.0	52.2	53.1	52.6	60.4
Volume of shipments	37.3	31.5	42.3	57.1	57.1	56.3	56.0	53.7
Volume of new orders	39.8	29.7	42.3	54.1	58.2	58.5	58.7	60.4
Backlog of orders	17.0	18.9	23.9	25.0	30.0	33.8	32.9	31.5
Inventories								
Materials	−15.8	−11.3	−10.9	−5.3	−10.5	−1.3	−16.4	−3.8
Plants' finished goods	−11.5	−18.1	−24.5	−19.8	−11.6	−15.6	−8.3	−6.0
Number of employees	10.2	8.0	10.7	24.0	18.7	22.0	25.0	20.8
Average employee workweek	4.4	−8.0	6.3	11.1	9.9	6.3	11.8	9.6
Prices received for finished products	14.4	24.5	26.6	28.9	27.5	21.5	21.9	49.0
Prices paid for raw materials	36.3	41.1	33.9	27.6	29.7	36.3	31.5	41.2
Capital expenditures	19.8	26.2	23.4	24.7	12.4	23.0	15.7	27.5
New orders for exports	24.6	20.8	18.9	27.3	33.7	28.0	31.3	30.0
Supplier delivery time[a]	−12.4	−5.6	−10.3	−4.2	−16.3	7.4	11.0	7.4

[a] *Percent slower minus percent faster.*

from the Atlanta survey for month-ago and six-months-ahead comparisons during the December 1991-July 1992 period.

By showing to what degree the number of plants with gains offsets those indicating worsening conditions, diffusion indexes indicate the direction of general trends for particular facets of manufacturing. The higher the index number in absolute terms (the closer the index is to 100 or to −100), the greater is the similarity of change among the responding firms. As indicated above, the closer index values are to 100, the more prevalent gains are among respondents; the closer values are to −100, the more prevalent declines are. An index level of zero indicates that the number of plants expanding and those contracting is evenly balanced.

Conclusion

The *Survey of Southeastern Manufacturing Conditions* has proved useful to the Atlanta Fed in monitoring regional economic developments. Over time the accumulated survey data can be evaluated statistically along with other indicators to provide greater depth of information about both the regional and national economies. In the longer run, because its data base includes background information on plants for a large number of factors not available in other series, the survey may provide a basis for research on manufacturing's cyclical behavior.[8] The data aggregated by various background factors could provide valuable insight into the long-run behavior of manufacturing plants in the Southeast.

Notes

1. The Beigebook is sent to Congress and the Federal Reserve Board of Governors and is made available to the public through the news media. It is released about a week before meetings of the FOMC, which meets eight times a year.
2. The Sixth Federal Reserve District encompasses Alabama, Florida, Georgia, the southern halves of Louisiana and Mississippi, and the eastern two-thirds of Tennessee.
3. Manufacturers of tobacco products and leather are not represented because of their very small share in the region's manufacturing output.

   Industrial classification refers to the grouping of reporting establishments on the basis of their major product or activity as determined by the establishments' percentage of total sales or receipts. The Atlanta Fed's manufacturing survey data are currently classified in accordance with the *Standard Industrial Classification Manual* (Office of Management and Budget 1987). The SIC codes indicate the level of aggregation of data. A two-digit code signifies a broader level of aggregation than, say, a five- or six-digit classification.
4. The Census of Manufactures takes place only every five years; 1987 data are the most recent. When data for the 1992 Census of Manufactures become available, they will be used as a basis for stratifying the Atlanta Fed survey.
5. This method of measuring manufacturing production by industry is similar to that used by the Federal Reserve Board's national index of industrial production (see Board of Governors 1986). This procedure contrasts with counting physical units coming off an assembly line. Output for a given industry includes the value of the inputs. Counting output for all industries would involve double counting and overstate the contribution of manufacturing to overall output.
6. The Atlanta Fed manufacturing survey's diffusion index is analogous to the fairly well-known diffusion index of employment change published by the U.S. Bureau of Labor Statistics, which measures the percentage of industries that posted increases in employment over specified time spans such as monthly and six-month periods. (For example, see U.S. Department of Labor 1992.)
7. The Bureau of Labor Statistics (BLS) and the National Association of Purchasing Management (NAPM), for example, calculate their indexes by determining the percentage of positive responses and adding to this one-half of the no-change responses. This method produces diffusion index ranges from 0 to 100, with 50 being the break-even point. Values above 50 indicate that positive responses outnumber negative responses; values below 50 indicate the opposite.

   In both the Atlanta Fed and BLS/NAPM methodologies, respondents have only three choices for answers (decrease, no change, increase), so a diffusion index based on a combination of any two answers provides the same information as any other combination of the answers. (Any combination of two out of three responses will yield the same information as long as the combinations are linear in construction. Essentially, three variables are set equal to a constant [100 percent] and two of the variables are known.) Using either methodology, one could construct an index based on negative and no-change responses that would have the same relative movement.
8. This background information includes industrial classification size of firm by number of employees, geographic location (state), union or nonunion status, domestic or foreign ownership, and vintage of plant according to the date of plant construction or last major capital improvement.

References

Bell, John, and Theodore Crone. "Charting the Course of the Economy: What Can Local Manufacturers Tell Us?" Federal Reserve Bank of Philadelphia *Business Review* (July/August 1986): 3-16.

Board of Governors of the Federal Reserve System. *Industrial Production, with a Description of the Methodology*. Washington, D.C., 1986.

Chmura, Christine. "New Survey Monitors District Manufacturing Activity." Federal Reserve Bank of Richmond *Cross Sections* (Winter 1987/88): 9-11.

Federal Reserve System. *Current Economic Conditions by Federal Reserve District*. Published eight times annually.

Harris, Ethan. "Tracking the Economy with the Purchasing Managers' Index." Federal Reserve Bank of New York *Quarterly Review* 16 (Autumn 1991): 62.

National Association of Purchasing Management. *The Report on Business: Information Kit*. Tempe, Ariz.: NAPM Information Center, 1990.

Office of Management and Budget. *Standard Industrial Classification Manual*. Springfield, Va.: National Technical Information Service, 1987.

U.S. Department of Labor. Bureau of Labor Statistics. "Diffusion Indexes of Employment Change, Seasonally Adjusted" (Table B-6). *Employment Situation*. Washington, D.C., August 1992.

Review Essay

The Future of Banking

by James L. Pierce.

New Haven, Conn.: Yale University Press, 1991.
163 pages. $25.00.

Aruna Srinivasan

Must U.S. banks forever be regulated? Has regulation added to banks' problems? Should banks be regulated in the same manner as in the past, perhaps with some minor modifications? Surely the financial system is much different now than when the basic public policies toward banks were created in the 1930s. The debate about reform and redesign of these policies often fails to recognize such changes. Instead, discussions assume that the economy's credit and liquidity services will continue to be performed by banks, as they have been historically. An exception to this trend is *The Future of Banking*, which proposes fundamental changes in banking powers, deposit insurance, and bank regulation. The Twentieth Century Fund, a research foundation that analyzes economic, political, and social issues, asked James L. Pierce to write this book to help clarify the issues in the bank reform debate and thus aid policymakers in making better-informed judgments about how to cure the industry's problems. Pierce, who teaches economics at the University of California at Berkeley, was also asked to offer his own solutions to banking's difficulties.

Pierce's basic analysis hinges on a few key points: Because the banking industry is critical to the stability of the nation's financial system, regulation is essential to protect it. However, what has worked in the past is no longer effective. New regulations should be consistent with economic forces and should not isolate and protect banks from competitive realities. Although banks at one time provided essential financial services that were not available elsewhere, now most of these services are not unique. Conversely, banks have been barred from certain lines of business, such as securities, that today seem a logical and beneficial fit with other bank services.

The reviewer is an economist in the financial section of the Atlanta Fed's research department.
She thanks Larry Wall for helpful discussions.

Pierce advocates regulations based on the type of financial services provided regardless of the type of firm that offers these services. Under this scheme any firm offering a checking account—whether a bank, securities firm, or some other entity—would be regulated in the same way for that particular service. Pierce would extend the safety net afforded by federal deposit insurance and Federal Reserve discount window lending only to firms offering monetary services, leaving the nonmonetary services of financial service companies, with uninsured deposits, at the mercy of the free market in gathering and lending funds.

The changes Pierce outlines are far more sweeping than the provisions contained in the Federal Deposit Insurance Corporation Improvement Act of 1991 (FDICIA), but his plan is one of a number of alternative ideas that seek to limit bank regulation and deposit insurance to firms that offer demand deposits and invest these deposits in "high-quality" assets, such as Treasury securities.[1]

In contrast to Pierce's proposal, FDICIA attempts deposit insurance reform within the framework of the existing regulatory structure. FDICIA's approach assumes that problems in the banking system stem from regulatory mismanagement and lack of market discipline in the deposit insurance system. The law relies on prompt, capital-based corrective action, risk-based deposit insurance, and least-cost resolution to minimize losses to the deposit insurance fund. FDICIA also curtails too big to fail by limiting the FDIC's ability to cover uninsured depositors and restricting the use of the Federal Reserve discount window to prop up weak banks.[2]

This essay suggests that while the new banking structure proposed by Pierce falls short of the ideal, FDICIA may pose greater risks to the system because it significantly increases the regulatory burden on banks without affording them the protection from competition they have enjoyed in the past. While FDICIA appropriately seeks to limit moral hazard and introduce market discipline, it also risks creating future crises by not recognizing the contraction in banks' traditional role.

What Services Do Banks Perform?

To understand current banking issues, a logical starting point is to consider the "special" nature of services banks provide. Pierce believes that the banking industry developed in response to two economic needs. The first he defines as a monetary one, whereby bankers protect "money," account for its ownership, and facilitate its use in settling economic transactions. In this original role, banks did not provide a return to the holders of funds because they were really providing liquidity services for which they were entitled to receive compensation.

In their second role—making loans or enhancing credit—banks evolved into the most efficient institutions for taking funds from surplus units (depositors) and channeling these funds to deficit units (borrowers). Recent theoretical work on the economic functions of banks suggests that banks provide two major credit services.[3] First, banks provide information services that aid individual investors. For example, by purchasing bank liabilities, investors can avoid both the costly duplication of effort each would make in researching and analyzing credit risks and opportunities while avoiding the possibility that other investors could reap the benefits of their analysis without incurring the costs (free riding). The second credit service that banks provide is monitoring firms' managements at a lower cost than individual debtholders can.

Pierce appears to suggest that banks' credit and liquidity services were combined by historical accident, and he does not explain why bank debt should be of rather short maturity (for example, demand deposits). This omission is significant because the rationale for traditional bank regulation rests on the idea that banks historically have been firms that combine nonmarketable assets with demandable debt liabilities.

Recent papers by Charles W. Calomiris and Charles M. Kahn (1991) and Mark J. Flannery (1992) offer several reasons that banks have historically engaged in maturity mismatching. Their results suggest that this mismatching provided important economic benefits. Calomiris and Kahn show that demandable debt is an important part of an incentive scheme for disciplining bankers. Demandable deposits permit depositors to vote with their feet; they withdraw funds when they lose confidence in banks. Without the ability to make early withdrawals, depositors would have little incentive to monitor banks. Flannery demonstrates that maturity mismatching may be optimal in an unregulated environment. He evaluates optimal means of financing a portfolio of bank-type loans and shows that uninsured banking firms face asset substitution and investment problems, which are best addressed by shortening debt maturity.

Financing nonmarketable assets with demandable liabilities is problematic, however, in that it exposes banks to depositor runs. To address this problem

governments have chosen to intervene by insuring deposits and regulating banks.

Yesterday's Solutions

In response to a particularly devastating failure of the banking system during the Depression, the New Deal banking reforms of the 1930s were designed to resolve the problem of depositor runs. The most important of these reforms related to deposit insurance and the safety net. The Federal Reserve was to act as lender of last resort, providing liquidity to banks when depositors wanted their money and thus limiting systemic risk. The FDIC was to insure deposits to protect individuals' wealth and assure them of the banking system's safety.

Other notable financial reforms were enacted during this period. The Glass-Steagall Act, which separated commercial banking from investment banking and prohibited commercial ownership, restricted banks' ability to compete with each other and with other financial institutions. Pierce argues that restrictions on granting of new bank charters, combined with limitations on branching, granted banks a monopoly of sorts in their limited territories. Congress also imposed ceilings on deposit rates in an effort to aid bank profitability. According to Pierce, all of these reforms were designed to prevent banks from failing and needing government assistance in the long run.

The Roots of Banking's Current Problems

The banking system put in place during the New Deal worked unusually well, Pierce concedes, until the 1960s, when the interaction of an inflexible regulatory system and changing banking structure led to problems in the financial services industry. At that time, changes in the structure of money markets and major improvements in technology allowed competition to flood into banking. Competitors, including the commercial paper market and money market mutual funds, offered less expensive, more flexible products that effectively challenged the profitability of the banking industry. The effects of these financial innovations on banking were so profound, Pierce contends, that understanding them is essential to any meaningful discussion of public policy toward banks.

Pierce charges both regulators and bankers with mismanaging the financial system during the 1960s and 1970s. Regulators, he argues, refused to adjust their actions in response to changes in the financial system and thus hindered banks from developing products that would compete with the new financial instruments from nonbank sources until the competitors were well entrenched in the market. Banks also made substantial errors in judgment, according to Pierce. Improved technology and access to money markets implied that large banks could grow rapidly by managing their liabilities, but this growth was accompanied by increasing leverage and declining asset quality, exposing banks to greater risks from loan losses and other sources. The substantial increase in leverage was allowed chiefly because bank deposits were protected by government insurance.

FDICIA's Objectives

Problems that arose in the banking industry during the 1960s and 1970s have become more pervasive. Since the early 1970s depository institutions have been plagued by persistent financial difficulties. The problems of the thrift industry and loan losses at large banks have received widespread publicity. Even as large banks in New England and Texas have required massive amounts of federal assistance, the bank insurance fund has sustained losses, raising the possibility of taxpayer bailouts.

The hiatus in legislative activity relating to federal deposit insurance finally ended, however, with the passage of FDICIA. The act represents, among other things, the first attempt to modify the deposit insurance system created during the 1930s. Under FDICIA capital becomes the centerpiece of bank regulation. The act requires banking regulators to divide all banks into five categories according to their capital ratios and specifies actions, increasing in severity as a bank's capital ratio moves down the scale, that regulators must take for banks falling in each category.[4] These provisions are intended to ensure prompt regulatory action when a bank first experiences difficulty and "early closure" when those problems (as measured by capital ratios) become severe. Regulators are also required to revise risk-based capital standards to take into account additional measures of risk such as interest rate risk, establish a system of risk-based deposit insurance premiums that would presumably rely heavily on capital ratios, and limit the

ability of banks in lower capital categories to acquire brokered deposits.

FDICIA requires the FDIC to establish a risk-based deposit insurance scheme by January 1994 and to study the feasibility of establishing a private reinsurance system. FDICIA also addresses the long-standing problem created by discount window lending to troubled institutions. It prohibits the Federal Reserve from making loans to an undercapitalized institution for more than 60 out of 120 days.

In addition to capital-focused mandates, FDICIA requires regulators to prescribe for all insured institutions operational and managerial standards covering such items as executive compensation restrictions, internal control standards, underwriting standards, interest rate exposure, and asset growth. The act also significantly increases banks' reporting requirements and includes new consumer protection provisions relating to Truth in Savings and CRA disclosures.

While FDICIA makes key changes in deposit insurance, it does so within the framework of the existing regulatory system.[5] FDICIA does not lower the $100,000 limit on deposit insurance coverage, it does not categorically prohibit regulators from protecting uninsured depositors at large institutions, nor does it privatize deposit insurance. FDICIA is a "narrow" piece of legislation in that it does not allow interstate branching, repeal Glass-Steagall provisions, or permit ownership of banks by commercial firms.

Pierce's Solutions

While Pierce acknowledges that there is scope for further reform within the existing regulatory structure, he believes that reforms based on this approach are likely to be effective for only a short time because competitive forces will continue to erode the traditional role of banks. Federal deposit insurance reform in particular, he argues, will not work without radical structural changes. Pierce discusses several reforms that could pave the way for more fundamental changes he has in mind.

Pierce wants to restore market discipline to banking by reducing or eliminating insurance on most categories of deposits. He recommends that banks be required to issue subordinated debt that would force added monitoring of banks by debtholders. Under Pierce's plan coinsurance would cover a given percentage of deposit balances in excess of the $100,000 statutory limit and would serve to remind large depositors that they need to apply discipline in selecting banks. The FDIC and uninsured depositors would share in the cost of bank failures.

Like other reformers, Pierce advocates restricting the maximum amount of insurance protection available to depositors. He suggests limiting the number of banks at which a depositor may carry fully insured deposits. Deposits spread among several banks would be coinsured rather than fully insured. Pierce contends that these proposals represent a slow retreat from total FDIC coverage and would reintroduce market discipline in stages.

Pierce's most distinctive reform would create monetary service companies as separately capitalized companies within banks or financial services firms. Monetary service companies would isolate, insure, and protect monetary functions. Only monetary companies would offer federally insured deposits and provide payments services. These companies would hold only money market instruments such as Treasury bills, commercial paper, and other short-term, liquid, highly rated instruments typical of money market mutual fund assets. Such a company could not lend to its owners under any circumstances and would be completely insulated from its parent company's liabilities. All other activities currently thought of as banking would become uninsured and unprotected by access to the safety net.

A separate financial services company, also under the umbrella of a larger banking or financial services firm, would provide nonmonetary services such as time and savings deposits and all lending functions. Deposits in a financial services company would be uninsured. Thus, if the company failed and investors lost their money, it would be solely a private-sector predicament. Depositors and other creditors of such companies would be forced to look more closely at where they placed their funds, Pierce argues, and this scrutiny would impose a healthy discipline on the banking industry.

Under Pierce's plan government regulation and insurance would be confined to only about 40 percent of what is currently defined as banks' liabilities. Commercial banking as it has been traditionally viewed, Pierce points out, has become a smaller component of the overall financial system. He estimates that currently no more than 37 percent of U.S. banks' sources of funds are made up of deposits payable on demand. Most liabilities are unrelated to the payments system, he contends. Only 18 percent of bank assets are devoted to commercial and industrial lending—the type of lending that historically has made banks special. Small

business lending, the one area in which banks have special expertise, is probably less than 10 percent of what the banking industry does. Pierce questions the necessity of regulating the other 90 percent of bank activities to ensure provision of this credit.

Although several of Pierce's bank reform proposals—including his narrow bank plan or the idea of breaking up the existing financial system—are not new or unique, they also are no longer discussed only in academic circles. FDICIA's reforms offer a framework against which Pierce's or any alternative proposal can be judged.

Problems Posed by FDICIA

Capital-based regulation, a key element of FDICIA, draws heavy criticism from Pierce. Such regulations cannot be implemented effectively, he believes, because they invite circumvention. For example, banks circumvented capital regulation during the 1980s by holding riskier assets and by increasing their off-balance-sheet activities. Regulators responded by developing risk-based capital standards, which weight various asset categories by their degree of risk. Risk-based capital standards are complex and irrational, Pierce asserts, and they distort banks' portfolio choices because they treat risks as additive. He fears that FDICIA's capital-based regulations will prompt a similar cycle of banks' finding new ways of circumventing the regulations and regulators' responding with restrictions on the new techniques for circumvention.

Pierce rightly raises concerns about FDICIA's early closure provisions. He points out that timely recognition of declines in asset values is crucial to implementing these provisions effectively. While FDICIA calls for a review of accounting rules for banks and thrifts, it does not mandate the use of market value accounting. Allen N. Berger, Kathleen Kuester King, and James M. O'Brien (1991) discuss some of the conceptual, measurement, and verification issues associated with implementing market value accounting that arise because of banks' roles in solving information and monitoring problems. Their study suggests applying market value principles to cases in which they are most feasible (for example, for traded securities and securitized loans) and using a statistical procedure for situations in which market value accounting is problematic (to correct for changes in credit quality, for instance). Even if market values could somehow be determined, Pierce questions the benefits of using market value accounting to close institutions that are viable in the longer run.

Pierce believes that major practical problems hinder implementation of risk-based deposit insurance. In his view, pegging deposit insurance assessments to risk "ascribes to the regulators an ability to identify and measure risk that they do not possess." Market discipline imposed by large depositors and capital requirements would be more effective, he argues.

Many in the banking industry have raised concerns not only about FDICIA's deposit insurance provisions but also about other elements of the legislation they consider highly onerous, especially those that require regulators to scrutinize more carefully the actual day-to-day business of financial institutions. These concerns stem partly from the fact that FDICIA maintains or strengthens existing consumer and social responsibility requirements on banks without liberalizing existing banking laws such as the Glass-Steagall Act and restrictions on nationwide branching. These additional requirements amount to a new tax on banks, while banks remain constrained in their ability to find more efficient ways to serve their customers. In Pierce's view, tightening the intensity and scope of bank regulation as FDICIA has done will only push more banking activities outside of conventional banks, putting continued pressure on banks to contract, especially in the traditional loan-making business.

Problems with Pierce's Solutions

Proposals like the fundamental restructuring of the financial system Pierce advocates require careful elaboration. Pierce anticipates a number, but certainly not all, of the "what-ifs."

The biggest problem with Pierce's solution is its extremity. It would abolish traditional banks, even those that remain viable, and would break up a number of existing financial institutions.

Pierce examines the impact his narrow bank ideas would have on lending behavior and credit availability in the financial system generally. Mortgage and consumer lenders would tend to accelerate their trend toward securitization. Commercial loans that are not easily securitized would have to be held on financial services companies' balance sheets, funded by uninsured deposits. As a result, interest rates on such loans would almost certainly increase significantly. Pierce does not concede this point, even though he argues that

current insurance on time deposits acts as a subsidy to commercial borrowers.

Another concern raised by Pierce's system is that it might facilitate the diversion of credit from smaller to larger businesses. Larger financial organizations could choose to engage only in deposit taking in various regions of the country without offering lending services through their financial services holding companies. If these diversified firms were to attract deposits away from locally based nondiversified banks, smaller businesses that now rely on local institutions could find less credit available.

While narrow bank proposals like Pierce's would separate illiquidity risk from the payments system, financial system stability would remain an important policy concern. Studies by Calomiris and Kahn (1991) and Flannery (1992) suggest that maturity mismatching in the existing system provides important market benefits and that the economy may be no more stable with maturity-matched banks because new firms that move in to fill the vacuum left by the demise of traditional banks may inherit the problem of depositor runs. Runs on financial services companies would have the same impact on the system that runs on banks had more than fifty years ago. Flannery argues that, as a result of political pressures, liquidation of financial services companies may take the form of government bailouts and thereby limit the benefits of narrow bank reform.

Pierce believes that many small banks would survive his restructuring proposal even though they might have trouble attracting uninsured deposits. Small banks that have relied on deposit insurance to stay alive would probably disappear; the more profitable small banks might be able to survive by securitizing and servicing the loans they originate.

Like any system of reforms, Pierce's proposal has several weaknesses. Ultimately, however, the most important issue is whether his restructuring scheme raises problems more severe than would occur under alternative frameworks such as FDICIA.

Policy Alternatives

What are the alternatives to the bank reform approaches discussed in this essay? One approach would be to weigh carefully all existing regulations and discard any that exceed the minimum regulation needed to prevent systemic risk and to protect the insurance fund. This strategy, which would place far greater reliance on market forces, probably has very little chance of being adopted.

Another solution would be to make some version of the narrow bank an option for any firm wishing to provide monetary services. This approach would permit traditional banks to continue operating in situations in which the costs of regulations are exceeded by the benefits and would allow for services that cannot be efficiently performed by banks (under current regulations) to be shifted to more efficient providers. Thus, monetary service providers would continue to operate under the safety net.

Overall, *The Future of Banking* is well worth reading. Pierce provides valuable insights into banking's problems and focuses the reader's attention on some fundamental issues relevant to the debate about the future of not only banking but of the entire financial services industry.

Notes

1. Pierce's monetary service companies are a variant of the core banks proposed by Bryan (1991) and the narrow banks proposed by Litan (1987).
2. See Carnell (1992) for a general discussion of FDICIA and Wall (1992) for a discussion of too big to fail and related provisions.
3. Gorton and Pennacchi (1990) summarize the major studies in this area.
4. See Carnell (1992) for a description of FDICIA's capital categories and rules governing institutions in those categories.
5. Congress required the FDIC to study the feasibility of authorizing insured depository institutions to offer both insured and uninsured deposit accounts, perhaps leaving the door ajar for more fundamental reform.

References

Berger, Allen N., Kathleen Kuester King, and James M. O'Brien. "The Limitations of Market Value Accounting and a More Realistic Alternative." *Journal of Banking and Finance* 15 (September 1991): 753-84.

Bryan, Lowell L. *Bankrupt: Restoring the Health and Profitability of Our Banking System.* New York: Harper Collins, 1991.

Calomiris, Charles W., and Charles M. Kahn. "The Role of Demandable Debt in Structuring Optimal Banking Arrangements." *American Economic Review* 81 (June 1991): 497-513.

Carnell, Richard Scott. "Implementing the FDIC Improvement Act of 1991." Paper presented at the Conference on Rebuilding Public Confidence through Financial Reform, The Ohio State University, Columbus, Ohio, June 1992.

Flannery, Mark J. "Debt Maturity and the Deadweight Cost of Leverage: Optimally Financed Banking Firms." University of Florida, unpublished paper, June 1992.

Gorton, Gary, and George Pennacchi. "Financial Innovation and the Provision of Liquidity Services." The Wharton School, unpublished paper, May 1990.

Litan, Robert E. *What Should Banks Do?* Washington D.C.: The Brookings Institution, 1987.

Wall, Larry D. "'Too Big To Fail' after FDICIA." Federal Reserve Bank of Atlanta unpublished paper, 1992.

FEDERAL RESERVE BANK OF ATLANTA

Public Affairs Department
104 Marietta Street, N.W.
Atlanta, Georgia 30303-2713
(404) 521-8020