Economic Growth Theory, Clustering, and the Rise of the South

G. Andrew Bernat, Jr.*

Abstract: This paper uses the relatively strong economic growth of the South over the past 20 to 30 years as the basis for a discussion of recent developments in economic growth theory. First, three facts about the South’s rise relative to other regions are discussed. This is followed by a discussion of the two leading macroeconomic explanations of economic growth, neoclassical growth theory and endogenous growth theory, showing how endogenous growth theory’s focus on innovation as the engine of growth, and on the externalities related to innovations, might be used to explain key regional characteristics of southern economic growth, such as clustering and uneven growth. The paper concludes with some observations about how the South’s future might unfold in light of these newer growth theories and briefly discusses how the resurgence in interest in regional economic growth presents excellent opportunities for regional science.

I. INTRODUCTION

According to one longtime observer of the South, “the South heads into the 21st century with the look and feel of prosperity...” (Autry et al. 1998, p 4). That the South can now be viewed as a prosperous, growing region is, I believe, one of the great success stories of recent economic history. By many measures, the southern economy has outperformed much of the rest of the nation over the past 20 to 30 years. Since 1969, jobs, per capita income, and population all grew faster in the South than in the nation as a whole. The natural question is, why has the South done so well relative to other regions?

This is an important question for a variety of reasons. It is certainly important for Southerners because knowledge of the factors that contributed to past growth is likely to improve our ability to predict and influence future growth. But its importance extends far beyond the South. For instance, the recent proliferation of studies of regional economic growth is due in large part to uncertainty about how increasing integration of international economies is likely to affect national economies. The U.S., with essentially no internal barriers to trade or to factor mobility, is seen as a laboratory for studying what is likely to happen as barriers to international cross-border flows of goods, services, and people, are reduced (Barro and Sala-i-Martin 1992).

In my talk today, I would like to use the rise of the South to discuss some
of the newer ways of looking at economic growth. In order to maintain consistency with previous literature on the South, I include in the South BEA's Southeast region plus two states from the Southwest region, Texas and Oklahoma. Arizona and New Mexico, the remaining two states in the Southwest region, are moved to the Rocky Mountain region. My talk is divided into three parts. First, I'll discuss three facts about regional growth and the South's rise relative to other regions. I'll then briefly discuss two competing macroeconomic explanations of this growth. Because the literature on economic growth is so vast, I won't attempt to be comprehensive.1 Instead, I will summarize what in my mind are a few of the key ideas in the current macroeconomic growth literature and show how these ideas may be useful to research in regional science. Specifically, I will try to show how recent growth theory's focus on innovation as the engine of growth, and on the externalities related to innovations, might be used to explain key regional characteristics of economic growth, such as clustering and uneven growth. I'll conclude with some observations about how the South's future might unfold in light of these newer theories and briefly discuss how the resurgence in popularity of regional economic growth in the economics profession at large presents excellent opportunities for regional science.

II. REGIONAL GROWTH AND THE RISE OF THE SOUTH

There are many ways to characterize regional economic growth in the U.S., but to me three facts stand out. First, regional economic growth is geographically highly concentrated. As I mentioned above, the South has experienced substantial economic growth over the past 30 or more years, with the number of jobs nearly doubling between 1969 and 1997. Rather than being spread evenly throughout the region, most of these 25 million new jobs were located in relatively few places. For example, eight areas made up of 11 counties, or less than 1 percent of the 1,366 southern counties, accounted for 25 percent of the South's net job growth while 42 counties accounted for 50 percent of net job growth. Furthermore, rather than evening out previous concentrations, this growth largely took place in existing concentrations of economic activity, as reflected by the fact that the 42 counties accounting for half of the net job growth are metropolitan.2

The second fact is that the South's economy has improved relative to other major regions. According to many standard indicators of economic well-being, the South is more similar to the rest of the nation than ever before. It was not so long ago that the South had the lowest wages and the lowest levels of per capita income in the nation. In 1969, the earliest year for which data on earnings per job are available, average earnings per job in the South was 85 percent of the U.S. average, 7 percentage points lower than in the Rocky Mountain region, the region with the second lowest average. But, by 1984, the South overtook both the Rocky Mountain and the Plains regions and remains above them today. Similarly, per

---

1 For recent reviews of the growth literature, see Temple (1999) and Durlauf and Quah (1998).
2 Also supporting the view that this growth has reinforced existing concentrations are the fact that 80 of the 100 counties with the highest job growth rates were metropolitan and the fact that the number of counties accounting for 50 percent of southern jobs fell from 74 in 1969 to 61 in 1996.
Economic Growth Theory, Clustering, and the Rise of the South

capita income in the South was, by a large margin, the lowest in the nation for many years. In 1958, for instance, per capita income in the South was only 78 percent of the national average. However, southern per capita income has risen steadily over the years to approximately 90 percent of the national average, and has exceeded per capita income in the Rocky Mountain region since the late 1980’s.

The third fact is that regional growth is not a steady, uniform process of convergence. While there is wide agreement that regional incomes in the U.S. have in general been converging over the years, not all regions appear to be converging. Overall, per capita income in three of the seven regions was farther from the U.S. average in 1997 than in 1958, closer in two, and about the same in the remaining two. Thus, while relative per capita income in the South has moved closer to that of the U.S. since 1958, average per capita income in the Rocky Mountain region has moved farther away, falling from 6 percent below the U.S. average in 1958 to 11 percent below in 1997. Divergence also occurred at the other end of the income distribution. Per capita income in New England rose from 9 percent above the national average in 1958 to 20 percent above in 1997. This is almost the same size increase in relative income recorded by the South but starting from above the U.S. average rather than below.

I’ll now turn to a discussion of how current theories of macroeconomic growth can explain the three facts discussed above.

III. TWO VIEWS OF REGIONAL GROWTH

The two major competing macroeconomic theories of economic growth are neoclassical growth theory and endogenous growth theory. Looking only at macroeconomic growth theories at a conference of regional scientists may seem a little unusual, but as Martin and Sunley (1998, p. 202) point out, the macroeconomic tendency to focus on aggregate, long-run phenomena can be a useful counterbalance to the tendency of regional scientists, geographers, and sociologists to focus on individual regions and localities. Thus, some of the insights from the macro growth literature may prove useful to regional scientists, just as some of the work by international trade theorists like Paul Krugman have proven useful.

A. NEOCLASSICAL GROWTH THEORY AND THE RISE OF THE SOUTH

The basic neoclassical model of economic growth asserts that economic growth can be modeled as a function of growth in capital, labor, and technology. From a regional perspective, this model has two key assumptions.

First, the marginal productivity of capital is assumed to be diminishing. In other words, an additional unit of capital increases output less than the previous unit. This assumption is important because it leads to one of the key predictions of the neoclassical model, namely, that incomes in different economies will tend to converge over time. The reasoning behind this prediction is fairly simple. Because the marginal product of capital is diminishing, the returns to an additional unit of capital investment are lower in regions with relatively large capital stocks than in

3Data on state per capita income is available as far back as 1929 but the earlier data may be less reliable.
regions with relatively small capital stocks, creating an incentive for capital to shift from capital-abundant regions to capital-scarce regions. As capital increases in the previously capital-scarce regions, the capital-to-labor ratio rises, leading to rising wages, with the reverse occurring in the capital-abundant regions, all other things equal. The effect on labor is similar, with labor tending to shift from relatively labor-abundant regions to labor-scarce regions. The net result of these shifts is an overall tendency for returns to both capital and labor to converge to the economy-wide average.

The second key assumption is that the level of technology is a function only of time and is unrelated to either capital or labor. The assumption of diminishing returns to capital means that the growth rate of output per worker declines in the long run (Grossman and Helpman 1994, p. 25; Martin and Sunley 1998, p. 207). However, one of the salient characteristics of economic growth over the past 100 years or more is the substantial growth in labor productivity (Hansen and Prescott 1998). In order to explain this growth in productivity, a term representing the level of technology is included along with capital and labor in the standard neoclassical production function.

From a neoclassical view of the world, the rise of the South is the result of a shift of capital from the capital-abundant, high-wage Rustbelt to the capital-scarce, low-wage South (Chinitz 1986). According to this explanation, wages rose in the South relative to the rest of the nation because capital increased faster in the South than did the labor supply, raising the capital-labor ratio and thereby raising the marginal product of labor. These shifts resulted in relative wage growth in the South. Put simply, the neoclassical view says that the South rose because it was initially so far behind and it will continue to rise until it has completely closed the gap.

The simplicity of this basic neoclassical growth model makes it very attractive, at least to economists searching for mathematically elegant equilibrium models of economic growth, but it falls short of an adequate explanation of southern growth in particular and regional growth in general. Of the three key facts about regional growth, the neoclassical approach is consistent only with the rise of the South. The other key facts, the divergent growth of some regions and geographic clustering, are difficult to reconcile with the model’s theoretical predictions.

Because the basic neoclassical model predicts only convergence, it simply cannot explain the movement of the New England and the Rocky Mountain regions away from the national average. In fact, comparisons of international economies indicate quite clearly that there is no tendency for poor nations as a whole to catch up to rich nations, as predicted by the neoclassical model.4

The failure of the basic version of the neoclassical approach to explain extended periods of divergent regional growth has not led to its abandonment. After all, growth theory is about long-term trends and it is possible that in a few more years we will see relative per capita income fall in the New England and the Mideast regions and rise in the Rocky Mountain region. Another possible expla-

4See, for example, Temple (1999).
nation is that convergence has been attained. While I'm skeptical of these explanations, it is nevertheless possible that, with a fair amount of tinkering, the neoclassical model can be made consistent with the divergent growth, perhaps even persistent divergence.

Much more serious, at least from the perspective of explaining regional growth, is the neoclassical model's inability to explain the spatial concentration of growth and the persistence of clusters of economic activity, both of which are fundamental characteristics of the regional landscape. In order for clustering to exist, increasing returns to some factors of production must also exist (at least over a range of values) but diminishing returns are fundamental to the neoclassical model. Consider urban growth. In order for cities to grow, as they clearly have in the South and elsewhere, capital and labor must be attracted to them. But cities, by their very nature, are locations with relatively high amounts of both human and physical capital. This seems to suggest quite strongly that the marginal productivity of capital is not declining, at least spatially (and for possibly a wide range of city sizes).

B. ENDOGENOUS GROWTH THEORY AND THE RISE OF THE SOUTH

The neoclassical model has many adherents but there is also widespread dissatisfaction with how poorly it seems to match the empirical record. While the neoclassical model predicts convergence, there is no evidence that poor nations, as a group, have been catching up to rich nations. In addition, many feel that the neoclassical model provides an unsatisfactory explanation of economic growth because of the exogenous nature of technological change, especially in light of substantial, if not steady, growth in output per worker. If technology is the engine of growth, then the production and use of innovations should play a prominent role in a theory of economic growth. The term innovation is used in the growth literature to mean any improvement in technology and includes new ways of producing existing products as well as new products themselves.

The dissatisfaction with the neoclassical model led to the development of endogenous growth theory. While there are a number of variants of endogenous growth theory, they all incorporate the notion that there are “knowledge spillovers” associated with capital investment. The term knowledge spillover refers to the spread of information about technology or markets from one firm to another. For example, suppose a firm develops an improved method of producing

---

5 Empirically determining when convergence has occurred is not a simple matter for a wide variety of reasons. For the view that convergence has been achieved among states see Evans and Karras (1996) and Carlino and Mills (1996).

6 I'm skeptical of these two explanations because the divergent growth of New England has been going on too long to dismiss it as a temporary aberration. In addition, New England is farther away from the national average in 1997 than the West was in 1958—clearly divergent behavior. A large number of studies have augmented the neoclassical model in various ways in attempts to make its predictions better fit the historical record. See, for example, Sala-i-Martin (1996) and Mankiw, Romer, and Weil (1992).

7 A similar argument holds with regard to labor. A strict interpretation of the neoclassical model appears to be at odds with the fact that people with high levels of human capital leave areas in which human capital is scarce and settle where human capital is abundant (Lucas 1988).
a particular product. A knowledge spillover occurs when other firms find out about the new method and use it to improve their production processes.

Knowledge spillovers are important to growth theories because of the nonrival nature of knowledge and ideas. The term nonrival refers to the fact that, unlike a new machine, a new idea that improves productivity in some way can be shared by everyone. The development of new knowledge and ideas thus benefits not just the developer of the new knowledge but society as a whole. In this way, knowledge spillovers impart an externality to capital investment, which in turn makes it possible to have a nondiminishing marginal product of capital. Without diminishing returns to capital, not only is geographic clustering possible but the prediction of economic convergence is no longer a necessary result.

Martin and Sunley (1998), in an excellent review of endogenous growth theory from a regional perspective, classify endogenous growth models into three groups: broad capital models, innovation models, and Schumpeterian innovation or neo-Schumpeterian models. These models differ primarily in the way the capital externality enters into the model. Broad capital models assume that investment in physical capital generates knowledge spillovers that raise the productivity of all capital in the economy (Romer 1994). Innovation models assume that knowledge spillovers are created not by investment in physical capital but rather by investment in human capital (Lucas 1988). Knowledge spillovers in these models affect the productivity of both human capital and physical capital.

Innovation in both of these types of models is, in a way, just a side effect or by-product of capital investment. In contrast, Schumpeterian innovation models assume that firms purposefully engage in innovative activities. Schumpeterian innovation models assume that imperfect competition exists, which makes it possible for firms to capture monopoly profits from their innovations. These profits are only temporary because of eventual knowledge spillovers, but the existence of profits, even if temporary, provides incentives for firms to devote resources to research and development (R&D) activities.

Dropping the assumption of perfect competition is clearly an improvement over the other endogenous growth models because doing so makes it easier to explain why many firms devote substantial resources to R&D. Dropping this assumption also opens other doors. By assuming imperfect competition, Schumpeterian innovation models are neatly aligned with the literature on what has come to be known as the New Economic Geography (NEG). This literature, associated with Krugman’s (1991) core-periphery model and Venables’ (1996) vertical-linkage model, shows how geographic clustering can occur even in the absence of knowledge spillovers.8

The logic behind the NEG is that the existence of Marshallian externalities leads to geographic clustering if production is subject to increasing returns to scale. The term Marshallian externalities refers to the fact that locating near input suppliers and near a diverse labor force can lead to geographic clustering of firms

8See Puga (1999), Ottaviano and Puga (1997), and Baldwin (1999) for reviews of this literature.
because production and distribution costs are lower for all the firms in the cluster. The connection between macro growth theories and the NEG is important because there are many industries for which it is difficult to argue that knowledge spillovers are important, but the NEG makes it clear that one can get clustering even without knowledge spillovers.

**Innovation, Clustering, and Regional Growth**

In order to see more clearly how innovations are related to geographic clustering, it is useful to distinguish between the innovations themselves and the activities that produce innovations. As I hope to show in the following discussion, innovations, while the key to long-term, global economic growth—the macro engine of growth—may or may not result in geographic clustering or regional growth. Innovative activity, on the other hand, is much more likely to result in geographic clustering.

Let's look first at how innovations contribute to clustering. According to endogenous growth theory, knowledge spillovers are the ultimate source of rising output per worker in the long run. Geographic clustering will result from these knowledge spillovers if distance is a factor in the diffusion of knowledge. In other words, if nearby firms adopt productivity-enhancing innovations before faraway firms, innovation will contribute to geographic clustering because early-adopters, who are located near the innovators, will be more competitive, and will therefore grow faster, than late-adopters. This type of diffusion is often referred to as the contagion model of diffusion because proximity is required for diffusion to occur, just as the spread of a contagious disease requires contact among people (Harrison, Kelley, and Gant 1996). As an example, suppose a firm develops a new way to move parts from inventory to an assembly line that substantially improves productivity. When employees of the innovating firm leave it to take jobs at other firms, they bring with them knowledge of this new process, thereby spreading its use from one firm to another.

By nature, innovations and associated knowledge spillovers are largely unobservable and are therefore difficult to measure and quantify. Nevertheless, their existence is widely accepted. Commonly cited examples of clustering based on knowledge spillovers are the computer and related establishments in the Silicon Valley of California and the Route 128 corridor in Massachusetts, the financial district in New York City, furniture manufacturers in Tupelo, Mississippi, and carpet manufacturers in Dalton, Georgia. Recent research on patents and R&D expenditures shows that both exhibit clear patterns of spatial clustering (Anselin, Varga, and Acs 1996; Feldman and Florida 1994; Jaffe, Trajtenberg, and Henderson 1993).

Let's turn now to the clustering effects of innovative activity. Innovative activity contributes to geographic clustering, apart from any clustering due to the innovations themselves, because of Marshallian externalities. We can distinguish between two types of innovative activity, which, for lack of better terms, I'll call

---

9 Named after Alfred Marshall, who first described, in the 1920’s, how the clustering of firms can create externalities that lead to further clustering.
R&D based and industrial district (ID) based. By R&D-based innovative activity, I refer to research and development that is conducted in laboratories, research centers, universities, and other large-scale facilities that are operated primarily for the purpose of developing new products or processes. This type of innovative activity is likely to have clustering effects primarily through Marshallian externalities because R&D facilities tend to require specialized labor and specialized inputs, which, as shown by the extensive work in regional science on agglomeration economies and, more recently, by the NEG, leads to clustering as firms attempt to minimize transport costs and take advantage of labor pooling effects.

Clustering effects attributable to the innovations themselves may be weak or nonexistent for R&D-based innovative activity. R&D facilities are likely to be operated by large, multiplant, possibly multinational firms, employing scientists and engineers who participate in professional activities, such as publishing in journals and attending professional meetings. Innovations produced in such clusters are thus likely to be rapidly disseminated over large distances either by way of the global scientific networks or through the possibly far-flung branch plants of the firms conducting the research and development. In either case, distance is unlikely to be a major barrier to the diffusion of innovations.

I've used the term “industrial district” to refer to the second type of innovative activity because firms in industrial districts appear to produce innovations in very different ways from R&D-based innovative activity. Industrial districts are “networks of mostly small, linked but generally loosely coupled, spatially-clustered manufacturing companies...with each firm specialized in one or a few phases of a complete project” (Harrison 1992, p. 471). The key characteristic of industrial districts that leads to geographic clustering is that the firms in an industrial district are closely linked in developing new products and new production processes. The purported resilience of industrial districts is attributed to their high rate of innovation and their flexibility in both adopting and developing new technologies and new products (Stöhr 1986). ID-based innovative activity is likely to have strong clustering effects through Marshallian externalities because small firms are likely to be highly specialized, acquiring specialized products and services from local suppliers rather than producing these themselves. Innovations produced by ID-based innovative activity are likely to have stronger clustering effects than the R&D-based innovative activity because knowledge spillovers among small and medium-sized firms are more likely to be informal, diffusing primarily according to the contagion model rather than diffusing through widely dispersed, global networks of researchers, as is the case for more formal R&D activity. In fact, because innovations are developed informally as part of the production process, they may tend to be undocumented, making it difficult to separate the effects of the innovative activity in industrial districts from the innovations they produce.

10For good descriptions of industrial districts, see Piore and Sabel (1984), Harrison (1992), Harrison, Kelley, and Gant (1996), and Scott (1992).
IV. PROSPECTS FOR FUTURE SOUTHERN GROWTH AND THE ROLE OF REGIONAL SCIENCE

How then does an endogenous growth model explain past growth of the South and what does it tell us about future growth? Unfortunately, a fully articulated macroeconomic endogenous growth model, much less a regional one, has yet to be developed and empirically tested, a fact that has led to some criticism of the endogenous growth approach (Pack 1994). Consequently, no simple endogenous growth explanation of southern growth is possible. However, the theory clearly points towards innovative activity and the growth and vitality of economic clusters for at least a part of the explanation. If regional growth is based on innovation and clustering, then we need to know more about why some clusters grow faster than others and why a region, and the clusters within a region, might be more innovative than others. That southern growth has occurred primarily in economic clusters is, I think, quite clear. Whether this is because some southern locations are more innovative, as an endogenous growth explanation would have it, or whether southern growth was due to factors more broadly regional in nature is an empirical question that can be pursued.

Looking to the future, whether or not southern growth will continue to be as geographically concentrated as in the past will depend to a large degree on future changes in transportation and information technology. While some argue that location will continue to be a critical factor in economic development (Porter 1995), there are strong indications that distance is becoming less and less of a barrier to the flow of goods and services and of ideas (Atkinson 1998). For instance, recent studies at both the regional (Hewings et al. 1998) and the national (Feenstra 1998) levels indicate that trade in intermediate goods and services has been expanding rapidly. Such hollowing-out, as Hewings et al. (1998) call it, suggests that transportation is becoming less of a factor in determining where firms purchase their inputs, weakening the clustering effects of vertical linkages among industries. Likewise, improvements in information and telecommunication technology appear to be weakening clustering forces. For example, the Internet appears to be weakening clustering forces as it makes proximity less important for a wide range of activities, such as retail shopping, banking, and stock brokering, that, in the not too distant past, were conducted almost exclusively in person.

I’ve argued that Schumpeterian models of economic growth provide a promising framework for modeling regional growth. They are attractive because they can accommodate a much wider range of outcomes—such as clustering and uneven growth—than neoclassical models. Unfortunately, the cost of the greater realism of the endogenous growth approach over the neoclassical approach is that a neat, concise explanation of regional growth may not be possible.

The rather undeveloped condition of endogenous growth theory presents excellent opportunities for regional science. As some of the leading economic growth theorists have observed, endogenous growth theory needs to incorporate research and ideas from other disciplines (Solow 1994, p. 53). Endogenous growth
theory tells us that the regions that grow the fastest are the regions that either innovate the most or that are the first to adopt innovations, but it tells us little about which regions these are likely to be. Regional scientists, however, have long realized that to understand why some regions innovate and grow better than others, we need to look beyond purely economic factors. For example, regional scientists have looked closely at how regional development is affected by factors such as entrepreneurship (Malecki 1994), trust and collaboration among local firms (Scott 1992; Harrison 1992; Markusen 1996), sense of place (Bolton 1992), and networking (Malecki and Tootle 1996). By synthesizing this large and diverse body of work with economic growth theory, regional science has an opportunity to make significant contributions not only to our understanding of growth in the South and other regions but also to our understanding of the general process of economic growth.

Finally, the new theories of economic growth give regional scientists the opportunity to have significant impacts on policy. Many research questions relating to regional growth, especially those that address the roles of innovation and clustering, will have an immediate impact on regional development policy because states and localities are rapidly embracing the concept of clustering as an economic development tool. Of particular importance for the South are what David Barkley (1998), in last year’s Presidential Address, called the “communities left behind.” He was not optimistic about the prospects of many of these communities, but the hope is that a better understanding of the issues I’ve raised today will lead to new policies that will keep at least some of these communities from being left behind in the future.

REFERENCES

Anselin, Luc, Atila Varga, and Zoltan Acs. “Local Geographic Spillovers Between University Research and High Technology Innovations.” Regional Research Institute, West Virginia University, Morgantown WV, October 1996.


Economic Growth Theory, Clustering, and the Rise of the South


