SPATIAL ECONOMIC ASPECTS OF THE ENVIRONMENT
AND ENVIRONMENTAL POLICY:
NEW DIRECTIONS FOR RESEARCH

by

Shelby Gerking
Department of Economics and Finance and University of Central Florida
PO Box 3985
Orlando, FL  32816-1400
e-mail: Shel Gerking@ucf.edu
Telephone: 407-823-3719
Fax: 307-766-5090

John List
University of Wyoming
University of Central Florida
PO Box 3985
Laramie, WY  82071-3985
Telephone: 307-766-4931
e-mail: Sgerking@uwyo.edu
Fax: 307-766-5090

ABSTRACT

This paper outlines major new research directions for research on spatial or interregional effects of environment and environmental policy. It attempts to define and characterize key research questions that lie at the intersection of environmental and resource economics and regional and urban economics. This intersection is of considerable importance because at least some factors of production are mobile, both domestically and worldwide. In consequence, information concerning the value of environmental attributes is revealed when people and firms relocate. Additionally, environmental policy changes can alter the interregional or international distributions of factor incomes, population, and production of goods and services.

For Presentation at the European Congress of the Regional Science Association
Vienna, Austria
August 1998
Although the first city smoke ordinances were enacted in 1855, widespread interest in environmental quality did not emerge until about thirty years ago. Passage of the National Environmental Policy Act in 1968, however, marked the beginning of an unprecedented increase in public awareness of environmental issues. Citizens groups formed to publicize specific environmental hazards and to press for corrective action. Government regulatory agencies, such as the U.S. Environmental Protection Agency, were established and then grew rapidly through the decade of the 1970s. Industries, particularly in the extractive, agricultural, and manufacturing sectors, began to make both environmental cleanup expenditures and investments in reducing future pollution flows. Academics began to think seriously about the consequences of environmental degradation, too. In economics, for example, a new field focusing on the environment and natural resources began to develop methods for determining the effects of environmental attributes and environmental regulation on economic activity.

This paper presents a survey of research on the spatial or interregional effects of the environment and environmental policy. Briefly, it attempts to define and characterize major research questions that lie at the intersection of environmental and resource economics and regional and urban economics. This intersection is of considerable importance because at least some factors of production are mobile, both domestically and worldwide. In consequence, information concerning the value of environmental attributes is revealed when economic agents relocate. Additionally, environmental policy changes can alter the interregional or international distributions of factor incomes, population, and production of goods and services. Quite often, these spatial effects either are not explicitly considered or are dramatically overweighted in the environmental policy development process.

The remainder of this survey is divided into four parts. Section 1 reviews the theory of how the spatial distribution of environmental amenities affects the interregional pattern of factor rewards. Section 2 discusses the relative role of amenities and economic opportunities in motivating migration. Section 3 treats valuation of environmental amenities using factor reward and migration data. Section 4 discusses the role of environmental policy in inducing industrial relocation. In all sections, recent literature is critically evaluated and future research directions are suggested.

1. **Theory**

   This section reviews theoretical contributions to the understanding of interregional economic aspects of the environment and environmental policy. Forster (1981) surveyed the early literature in these areas, most notable papers by Comolli (1977), Forster (1977) and Yohe (1979). In consequence, this discussion is not repeated here. Later papers that can trace their roots to Rosen’s (1974) seminal contribution on hedonic prices and implicit markets warrant greater emphasis because they either directly or indirectly inspired much of the empirical work treated in the next section. However, virtually all theoretical papers in this area, no matter what their origins, use general equilibrium models to address questions about how environmental amenities and/or environmental policies can affect the spatial distribution of income, production, and population.

   **a. The environment and interregional economic activity.**

   Roback (1982) further developed a general equilibrium approach initiated by Rosen (1979) to modeling effects of the environment on the spatial distribution of economic activity that continues to be influential today. As discussed in the next section, the perspective taken in her model continues to inspire controversy in migration research and serves as the basis for many recent empirical attempts to compute dollar values of amenities and urban quality of
life indices. The model allows environmental amenities to affect both utility of consumers and productivity of firms and its main contribution rests on showing how amenity values are determined through interaction of labor and land markets. Prior empirical literature had attempted to estimate values of environmental amenities using data on either wages (Nordhaus and Tobin 1972, Getz and Huang 1978, Rosen 1979, Henderson 1982) or land rents (Ridker and Henning 1967, Polinsky and Rubinfeld 1977), but data from both markets together.

Roback’s model posits a system of small regions in which each region employs an identical constant returns to scale production function to produce a single composite good (X) using labor (H), land (L), and the amenity (A). Workers, identical everywhere and spatially mobile at zero cost, consume X, use land for residential purposes, and enjoy the amenity. Land is the same in each region; in consequence, the only way to tell regions apart is by their differing endowments of A. The price of the composite good (P=1) is determined on a competitive world market and labor and land are paid their marginal value products (W for labor and R for land). Also, equilibrium has three characteristics that arise from the assumptions of competition, constant returns to scale in production, and perfect interregional labor mobility. First, worker utility levels are as high as possible and are the same in each region. Second, entrepreneurs earn zero profits and cannot achieve lower production costs by relocating. Third, labor and land are fully employed in each region.

This model can be used to demonstrate how variation in A between regions affects equilibrium factor rewards, output of the composite good, and population size (i.e., number of workers). Wages and land rents are uniquely determined via interaction of worker preferences and costs faced by entrepreneurs: Changes in the endowments of labor and land have no effect on these factor rewards. To more fully illustrate the mechanism by which factor rewards are determined, consider two regions (R_1 and R_2) that are endowed with differing amounts of the amenity (A_1 and A_2). Denote the indirect utility function of workers in R_i as (1) \( V_i = V(W_i, R_i, A_i) \) and the cost function of X producers in R_i as (2) \( C_i = C(W_i, R_i, A_i) \).

Let A_1 > A_2 and assume that the amenity reduces costs of producing X. Then, workers in the high amenity region (R_1) may be paid either higher or lower wages than in R_2, while R_1 land earns unambiguously higher rents than land in R_2. However, in spite of the ambiguous effect of amenity levels on wages, the wage-rental ratio always ends up lower in the high amenity region (R_1). Thus, the percentage increase in land rents is greater than any possible corresponding increase in wages.

Additionally, the spatial distribution of amenities affects the interregional distribution of X production and population of workers through the full employment conditions of the model. Roback (1982) does not discuss these conditions, although they necessarily arise from the assumptions of constant returns to scale in production and competition in all markets. Full employment conditions do come into play in her later work (Roback 1988). In any case, in the model at hand, they require that all land is used either for X production or for worker residences

(3) \( L = C_L X + kH \)

And that all workers are employed in X production

(4) \( H = C_W X \)

In equations (3) and (4), k denotes per worker utilization of land (and is a function of the W/R ratio), and \( C_L \) and \( C_W \) denote derivatives of the cost function which are interpreted as the per unit use of land and labor, respectively, in the production of X.

As previously discussed, supplies of L and H have no effect on equilibrium values of W and R; however, regions with differing amounts of A would be expected to produce different amounts of X and have different worker populations. For example, a region like R_1 above with relatively more A will have lower unit costs of producing X. Also, the lower
wage-rental ratio in R_1 means that: (1) X producers will economize on their use of land by utilizing more labor intensive production methods and (2) workers will have lower real incomes and consume less housing relative to the composite good. Thus, amenity levels affect X and H; however, the direction of effect is, in general, ambiguous.

A limitation of Roback’s analysis is that it treats only natural or nonproduced amenities. Gyourko and Tracy (1989) extend her model to include public services produced by local governments and taxes levied to finance them. In consequence, this extension allows for the possibility that local governments can choose public service levels, including those that may affect environmental quality. Workers and firms are assumed to be freely mobile between regions. Firms produce a single composite good under conditions of constant returns to scale that is sold on a world market at a fixed price (unity). Equilibrium wages and land rents are again determined through interaction of worker preferences and firm costs. The Gyourko and Tracy model, however, produces a broader set of comparative static results for changes in factor rewards in response not only to changes in natural amenities, but also to changes in government services and taxes.

As indicated at several points in the preceding discussion, the Roback and Gyourko and Tracy models have the property that, in equilibrium, unit costs and worker preferences uniquely determine factor rewards. This property also is exhibited in related models in which a nontraded good is added (Roback 1982) or two or more types of interregionally mobile workers are present (Roback 1988, Beeson 1991). The intuition here is that under the assumptions of competition in all markets, constant returns to scale in production, and zero mobility costs, n+1 interregional markets (one for goods and one for each type of labor) exactly determine n+1 factor rewards (see Ethier and Svensson 1986 for further discussion). This correspondence between the number of markets and the number of factors is in certain respects similar to standard textbook treatments of the 2x2 Heckscher-Ohlin (H-O) model in international trade theory in which factor rewards are determined by relative commodity output prices alone and factor endowments do not matter.

The Roback, Gyourko and Tracy, and related models, however, differ from the H-O model in an important respect. In the H-O model, commodity trade equalizes factor prices between countries (or regions). Thus, the issue of whether factors are mobile does not arise because there is no incentive to move. In the models reviewed here, workers must be freely mobile; otherwise, utility levels never would be equalized across regions. In fact, the assumption of free worker mobility is a key simplification that assists in breaking the link between factor endowments and equilibrium factor rewards. This point, which turns out to have implications for applying the model in a migration context, is explored more fully below.

To better appreciate the potential role of factor endowments and migration on the spatial distribution of factor rewards, consider three variants of the basic model. First, Blomquist, Berger, and Hoehn (1988) present a model in which a representative urban area in an interregional system is assumed to have two counties, each with a different level of amenities. (These authors also present a related model in which urban areas are composed of a central business district and surrounding residential zone (see Hoehn, Berger, and Blomquist 1987).) Their model allows for amenity variation both within and between urban areas. Additionally, production costs are subject to congestion or agglomeration effects and the model endogenously determines a city size variable incorporated directly into the production cost function. Other features of the model are similar to Roback’s (1982), except that production uses capital, but does not use land. Capital is spatially mobile and receives a return that is fixed by international markets.

In the Blomquist, Berger, and Hoehn model, equilibrium again occurs when neither workers nor firms can gain by relocating, so unit production costs always equal output price and utility levels are the same everywhere. The analysis focuses on how wages and land rents are affected by differences in amenity levels within an urban area. Cross-county effects
emerge because a change in amenities on one county results in a change in urban population, which in turn affects production costs of all firms located there. A curious feature of the Blomquist, Berger, and Hoehn model, however, is that city size is determined along with factor rewards through interaction of indirect utility and unit production cost equations. The fact that all factors must be fully employed (given the assumptions of competition in all markets and constant returns to scale production technology) is not utilized. Thus, although factor rewards depend on city size, city size is independent of the aggregate endowment of available labor.

As a second extension of the basic framework, assume that in the Blomquist, Berger and Hoehn model, the composite good had been produced using labor and interregionally immobile capital. Because capital is no longer mobile between regions, factor rewards to be determined (3) would outnumber interregional markets (2) (see Batra and Casas 1976, Ethier and Svensson 1986, and Jones 1971 for further details and interpretation). Factor endowments then would come into play (through full employment equations like (3) and (4)) in the determination of factor rewards. In this situation, wages and rents would still vary from one region to another, but this variation no longer would be entirely attributable to interregional differences in amenity levels. A possible implication of this outcome is that amenity values should be calculated net of factor endowment effects.

A third extension of the model arises when interregional labor mobility costs are introduced. For simplicity, assume that there is only one type of labor, although two or more types of labor easily can be considered. Mobility costs can drive a wedge between utility levels enjoyed in each region and, once present, these utility differences may persist over long periods of time. In this situation, equilibrium would occur when regional utility levels are not identical; but utility differences are not large enough to induce migration. Additionally, regional utility levels no longer would be exogenously fixed and, within limits determined by costs of migration, would be determined by, rather than a determinant of, wages, land rents, and other market prices in the interregional system. Consequently, worker preference and unit cost functions (see equations (1) and (2)) no longer would determine W and R uniquely and a link between factor endowments (via the full employment equations (3) and (4)) and these factor rewards would be established.

Three implications from this analysis are of interest in the present context. First, if labor mobility costs exist, the interregional pattern of wages and land rents would reflect both utility and amenity differences. This outcome complicates the problem of valuing amenities and computing quality of life indices; an issue considered in more depth later in this paper. Second, a reduction in mobility costs would tend to narrow interregional utility differences and would induce workers to relocate. Also, this worker migration, in turn, would alter the factor reward structure in each region. This outcome illustrates the general conclusion that when labor mobility costs are positive, factor rewards both determine and are determined by interregional factor movements. Third, worker migration can occur in response to a shift in tastes for amenities, but it also can reflect greater opportunities for consumption in the destination region that can arise simply because utility there is higher.

2. Amenities and Migration

The role of environmental amenities in interregional migration has inspired considerable debate. On the one hand, several papers by Graves and others for example, see (Graves 1979; Graves and Linneman 1979); argue that interregional labor markets are highly efficient and that any utility differences that arise are quickly arbitrated away through migration. As a result, observed migration occurs mainly, if not exclusively, in response to amenity demand changes that accompany increasing household incomes or lifecycle events such as retirement. On the other hand, the more traditional human capital view (Sjaastad 1962) is that utility differences between regions can be persistent because the process of labor market adjustment through migration is slow. This view accommodates the possibility that migration can be
related to increased demand for amenities; however, economic opportunities (i.e., not yet arbitraged utility differences) are generally thought to be the primary motivating factor. Evans (1990) and Hunt (1993) provide a comprehensive discussion of these two points of view in their surveys of equilibrium and disequilibrium approaches to migration modeling. In consequence, the objective of this section is mainly to examine recent attempts to elaborate on these perspectives.

Two articles, reviewed by Hunt (1993), represent empirical attempts to contrast the equilibrium and disequilibrium points of view. Greenwood and Hunt (1989) examine Graves’ contentions concerning the role of amenities in migration using two different data sets. They model migration as a function of both natural amenities and variables measuring economic opportunity. Coefficients of both types of variables turned out to be statistically significant; however, the economic opportunity variables, such as employment growth in a region, appeared to be more important in explaining migration. In a later paper, Greenwood, Hunt, Rickman, and Treyz (1991) consider more directly whether the interregional system in the U.S. is in equilibrium. In estimates using panel data for states, they again conclude that both economic opportunities and amenities play a significant role in influencing migration. Also, they measure the extent of interregional disequilibrium by comparing each state’s actual after tax real earnings with a corresponding earnings estimate that would yield no economic migration.

In a widely cited recent article, Mueser and Graves (1995) argue that migration is a reaction to disequilibrium and develop a dynamic model to explain the migration process. Central features of the model are that individuals face relocation costs as they consider migration to regions offering higher utility levels and that firms face adjustment costs in changing their levels of employment. These frictions prevent the spatial distribution of output and population from adjusting instantaneously and imply that the interregional system may not be in equilibrium. The model developed uses Roback’s (1982) definition of regional equilibrium and is used to explain how this steady-state position is attained when the system experiences no exogenous shocks for an extended period of time. Mueser and Graves contend that the mechanism depends on the relative speeds with which migration responds to utility differentials and employment growth responds to cost differentials.

The introduction of adjustment costs into the analysis, however, requires that Roback’s concept of long run equilibrium must be modified. As discussed at the conclusion of Section 1, migration only will occur as long as monetized utility differentials between regions exceed migration costs. Thus, interregional utility levels will never be equalized, factor endowments become a determinant of factor rewards, and long run equilibrium no longer can be characterized as before. Indifference curves in W,R space cannot be drawn on the assumption of a fixed utility level. Moreover, migration directly affects both a region’s utility level and its factor reward structure. Additionally, as Mueser and Graves (1995) point out, migration is most appropriately viewed as a response to interregional differences in utility. Spatial variation in amenities is a potentially important contributor to these differences; but interregional variation in opportunities for consumption of private goods may be important as well.

In their empirical analysis, Mueser and Graves estimate a reduced form regression to explain net migration between U.S. county aggregates in three decades; 1950-1960, 1960-1970, and 1970-1980. The net migration variable was estimated using cohort-survival methods. Explanatory variables included measures of industrial composition, amenities, settlement patterns, and demographic characteristics. Key findings are that migrants move to more desirable climates, avoid areas in which manufacturing and agriculture dominate the industrial structure, and avoid areas with a high percentage of blacks. Also, standardized coefficients for each group of variables indicated that the climate-related amenities dominated the industrial composition variables in explaining net migration rates. This procedure for comparing the importance of groups of variables, however, invokes the questionable
assumption that all relevant factors in a group can be enumerated and measured. Also, employment growth, a variable found to be important by Greenwood and Hunt (1989), was not considered as a possible regressor because it is endogenously determined in the Mueser and Graves dynamic model.

3. Interregional Amenity Valuation and Quality of Life Indices

The models outlined in Section 1 frequently have been used to support empirical estimates of the value of amenities and to compute indices of the quality of life. These studies have attracted wide attention within the economics profession as well as from government agencies and the business community. Economic development practitioners, for example, frequently pay close attention to quality of life rankings, particularly when their city or state is ranked at or near the top. This section reviews selected studies in this area with emphasis on those using multi-market, interregional data. It is divided into two subsections treating hedonic and revealed preference approaches.

a. Hedonic Studies.

Free mobility of labor has been a central assumption in hedonic estimates of amenity values and quality of life. This assumption implies that, in equilibrium, workers enjoy the same level of utility no matter where they live and that the desired estimates can be made from the spatial pattern of wages and land rents. A simple manipulation of the indirect utility function shows that the implicit price of the amenity \( P_A \) is

\[
P_A = k \left( \frac{dR}{dA} \right) - \left( \frac{dW}{dA} \right)
\]

In equation (5), \( k \) again denotes land used for consumption (residential) purposes per worker. The signs and magnitudes of the \( \frac{dR}{dA} \) and \( \frac{dW}{dA} \), of course, depend on the exact specification of the model as well as on whether \( A \) is an amenity in consumption or production or both. \( P_A \) is interpreted as the payment a worker must make for a small increase in amenity levels through changes in land rents and wages.

Roback (1982) empirically illustrates the use of equation (5) to compute values of amenities. Her results, which are based on wage data from the 1973 Current Population Survey and residential site price data from FHA-qualifying families, only are suggestive because the data on land prices were admittedly weak. Blomquist, Berger, and Hoehn (1988), then, conducted a broader empirical investigation based on housing expenditure and wage data available from the 1 in 1000 A Public-Use Sample of the 1980 Census, in which people are identified at the county level. Because their analysis used data on housing expenditure, rather than land prices, their model was reformulated to include housing in the utility function and to add a production function for housing. These alterations lead to valuing amenities based on a reformulation of equation (5) in which the right-hand-side becomes the derivative of the price of housing with respect to \( A \) weighted by the quantity of housing purchased less the wage effect, \( \frac{dW}{dA} \).

The analysis presents estimates of values for 16 different amenities such as precipitation, heating and cooling degree days, air quality, and extent of violent crime. A monetized quality of life index, is computed for 253 counties by summing the amounts of available amenities in each weighted by their full, implicit prices. Estimates of these prices were based on hedonic wage and housing expenditure equations, which included measures of the amenities as explanatory variables. Coefficients of the amenity variables generally were significantly different from zero and plausibly signed. Because these coefficients were constrained to be the same for each county, variation in the quality of life index between counties reflects only differences in amounts of amenities present. Pueblo, Colorado turned out to have the highest monetized quality of life ranking at $3289, followed by Norfolk, Virginia at $2106. St. Louis, Missouri had the lowest ranking at -$1857.
Gyourko and Tracy (1991) also estimate amenity values and constructed quality of life indices using 1980 Census of Population and Housing 1 in 1000 public use A sample. Estimates in this study pertain to 130 cities rather than counties. In keeping with their model (outlined above), a central focus of the analysis rests on the contribution of local fiscal variables, as contrasted with pure amenities, as well as the possibility that location rents may be at least partially capitalized into public sector wages. Thus, city traits accounted for in their wage and housing expenditure regressions emphasized these aspects and excluded some of the environmental measures used by Blomquist, Berger, and Hoehn. Additionally, Gyourko and Tracy included error terms in both their wage and housing expenditure regressions that reflected both individual-specific and city-specific components.

Coefficient estimates reported generally have expected signs. Estimated full implicit prices indicate that public services (i.e., health, education, and safety) are “goods”, while higher taxes generally are “bads,” although results for the corporate tax rate are anomalous. Also, both individual- and city-specific disturbance variances are significantly different from zero. The interpretation of this result, however, is uncertain; yet it holds considerable importance for calculation of quality of life indices. On the one hand, as Gyourko and Tracy indicate, the city-specific disturbance may reflect unmeasured housing quality or worker human capital, in which case it should not be included in the index. On the other hand, if it reflects omitted amenity or fiscal variables that would have been included in the regressions had appropriate data been available, then it should be included. Another possibility, not considered by these authors, is that the city-specific disturbance term may reflect utility differences between regions that arise because of imperfect worker mobility.

In any case, Gyourko and Tracy present quality of life rankings, with and without city-specific effects, calculated using the same approach adopted by Blomquist, Berger, and Hoehn (1988). For some cities, whether these group effects are included makes little difference; San Diego, CA ranked fourth under both approaches. However, without accounting for city effects, Boise, ID ranked 108th; whereas in the rankings with group effects it was second. Overall, the rank correlation between the two sets of estimates was 0.63. Pueblo, CO, which was ranked first by Blomquist, Berger, and Hoehn, ranked 66th when city effects are excluded and 96th with city effects are included. This example illustrates the fact that the quality of life rankings are imprecise and can change substantially with changes in the underlying amenity and public service variables as well as with changes in data sets and econometric methods employed. This issue is more fully discussed by Gyourko (1991).

b. Revealed Preference Studies

Kahn (1995, p.224) discusses four specific problems associated with the hedonic approach to calculating quality of life indices. First, all amenities and local public services must be exhaustively enumerated and observed. Second, unmeasured city-specific attributes may be correlated with included attributes, resulting in biased and inconsistent estimates. Third, included attributes may have a high degree of intercorrelation (recall that Blomquist, Berger, and Hoehn 1988 used 16 area-specific attributes in their wage and housing expenditure regressions). Fourth, returns to human capital characteristics may differ across regions. He attempts to avoid these problems by developing an alternative method for calculating quality of life indices that is based on revealed preference. This method is based on three assumptions: (1) migration costs are zero, (2) all agents in the economy have identical preferences defined over consumption and a set of city-specific attributes, and (3) all agents have equal skills. Then, if people in City A are observed to have lower consumption than people in City B, then City A must be the “nicer” of the two.

Kahn’s revealed preference method was implemented by ranking quality of life in five U.S. cities (New York, Houston, Chicago, San Francisco, and Los Angeles) using Census data for 1980 and 1990. Equations for wages and rental expenditures were estimated for each city. This approach allowed for inter-city variation in characteristic prices and the constant
term in each equation is interpreted as the net effect of all city-specific amenity and local fiscal variables. Estimates obtained were used to predict wage and rental expenditures for each sample member in each of the five cities. Surpluses then were calculated by subtracting predicted rental payments from predicted wages. Cities with the highest surpluses (Chicago and Houston) were judged to have a relatively low quality of life: The extra amount available for consumption is treated as compensation for lower levels of natural amenities and poorer local fiscal conditions.

Although the revealed preference approach has certain advantages over that used by Blomquist, Berger, and Hoehn (1988) and Gyourko and Tracy (1991), it has at least three limitations. First, as discussed by Rauch (1993) and noted by Kahn (1995), cities with larger amounts of human capital per worker may be more productive and have higher wages than other cities. Thus, a relatively large surplus for a city may reflect higher productivity rather than a lower quality of life. Second, the method assumes that all worker attributes are observed which, of course, may not be true. As Kahn (1995) points out, if workers with greater unobserved ability self-select into the nicest cities, relatively high surplus values may no longer correspond to lower quality of life. Third, workers are assumed to be perfectly mobile, so that utility is the same in all cities. However, if workers face relocation costs, utility levels will not be identical everywhere. In a given city, some workers may be earning positive quasi-rents, while the quasi-rents for others are negative. Kahn attempts to limit this potential problem by restricting his sample to renters, who would be more mobile than people who own homes, and by reporting the median of individual surpluses for a city. Nevertheless, utility differences may not “average out” and large surplus values still may reflect higher levels of utility, rather than a lower quality of life.

Herzog and Schlottmann (1993) implement a second revealed preference approach to estimating values of amenities by looking at migration data. Citing the Greenwood, Hunt, Rickman, and Treyz (1991) study, these authors contend that the assumption of equilibrium in the labor and land markets substantially weakens the applicability of the hedonic method to valuing amenities. They propose to correct this problem by estimating a migration (M) function in which the decision to migrate depends on wages, housing prices, and amenities. These variables are the arguments of their indirect utility function (which was specified quite similarly to equation (1) above). As a result, they reason that amenity values can be estimated from the total differential of the migration function. Setting dM=0 in the migration function implies that dV=0 in the indirect utility function, permitting monetary tradeoffs between wages, housing prices, and amenities to be calculated.

In their empirical analysis, Herzog and Schlottmann estimate their migration equation using the 1% Public Use B Sample of the 1980 Census. Amenities were proxied by urban population and population density, rather than enumerating them individually. They find that urban scale is a net amenity for cities with populations lower than 4.4 million people and a net disamenity for larger cities. These results were substantially different from implicit price calculations derived from estimates of a wage equation in which urban scale was found to be a net disamenity for cities of all sizes. However, the migration-based willingness to pay estimates appear to be deficient because setting dM=0 may not imply that dV=0. In fact, as the authors themselves point out in a footnote, dM=dV=0 only holds when there are no transactions costs associated with moving. Yet, as discussed in Section 1, equilibrium would emerge if this condition holds, thus ruling out further migration. Interestingly, the presumption of equilibrium was exactly what the authors sought to avoid by constructing the migration-based willingness to pay estimates.

Cragg and Kahn (1997) also use migration data to estimate the value of climate and quality of life. Their study is base on data from the 1% PUMS samples from the 1990 Census of Population and Housing. The idea here is to determine the extent to which migrants pay for amenities through changes in wages and rents when they move from one state to another. The sample consists entirely of migrants; thus, the decision to move is not considered.
Rather, a conditional multinomial logit model is estimated to explain which of the lower 48 U.S. states was chosen when migrants relocated. In this model, migrants are assumed to maximize utility by maximizing consumption of market goods and natural amenities. The value of market goods consumption was measured by the difference between after tax income and the rental price of a four-room home.

Migrants were assumed to know what their consumption bundle would be in each their possible 47 destination states. Thus, equations were estimated for wages, weeks worked, and home rental prices and consumption values were imputed for each migrant in each state. Wage equation estimates were significantly affected by whether returns to worker characteristics were constrained to be equal across states. In any case, the implied consumption values then were used to estimate parameters of the conditional multinomial logit model. These parameters, in turn, were used to make calculations of willingness to pay for environmental amenities. In particular, the authors ask (p.277), “if an environmental good’s quantity increased by one standard deviation, how much private consumption could we take away from that person such that his probability of moving to a given state would remain unchanged?”

The resulting estimates showed that: (1) the most highly valued environmental attribute was a higher average February temperature and (2) older people and non-college graduates valued environmental attributes more highly than did younger people and college graduates. Also, willingness to pay estimates for individual amenities exceeded corresponding estimates of actual payments computed from the hedonic wage and rental equations. State-by-state quality of life rankings using the migration approach placed warm winter states (Florida, Arizona, and California) at the top. The hedonic approach, which placed greater weight on cooler summers than warmer winters, ranked California 25th and moved Arkansas and Oklahoma to 3rd and 4th, respectively.

4. Environmental Regulation and Firm Location

The goal of this section is to analyze the empirical evidence regarding the relationship between environmental regulation and economic activity. However, given Tannenwald (1997), Levinson (1996a), Olewiler (1994), and Gunther (1991) have recently provided a thorough literature review in this area, the focus of this section will be on potential explanations of certain empirical findings. This section is divided into three parts: (1) a theoretical model of firm location, (2) a summary of empirical findings, and (3) potential avenues for future research.

a. Models of Firm Location

While many theoretical arguments have been advanced to examine the spatial patterns of firm-level investment, industrial organization theories have become the most prevalent due to their intuitive appeal. Dunning (1988) integrates these market structure theories and argues that investment can be explained by the interaction of three factors: (1) internalization advantages; (2) firm-specific advantages; and (3) locational advantages. Although internalization and firm-specific advantages are important in the decision process, recent theoretical models have focused on locational advantages. Markusen’s (1984) model suggests that firms invest to take advantage of their ownership of an input whose productivity remains the same given multi-plant operations. Extending the Markusen (1984) model, Markusen et al. (1993) use a two-stage model to show that plant location and the resulting market structure are endogenous to environmental policies adopted by different regions. Because a pollution tax (or a regulatory requirement to use high quality inputs) increases the marginal cost of production, firm profits change in the second-stage, leading firms to alter their location decisions in the first stage.

As an extension to the Markusen et al. (1993) model, consider a model that includes both site-specific fixed and marginal production costs. Alternative locations are assumed to have specific natural and/or policy-based features. Natural location features include proximity...
to markets and natural resources while policy-based features can arise due to a myriad of reasons, including tax and public expenditure policies, labor standards, and environmental regulations. To make this point clear, consider the investment decision of a firm that has decided to locate a plant in one of K locations. Denote the firm’s constant marginal production cost as $m_j$, where $j=1,...,K$, varies by location because of natural characteristics and/or policy differences. Because the firm also has to acquire a building permit, zoning approvals, and in the case of a polluting firm, pollution abatement equipment, site-specific start-up costs, $g_j$, must also be incurred. From duality theory, the firm’s problem becomes one of cost minimization:

\[
\text{Min } C = f(\bar{q}, m_j, g_j)
\]

where $j = 1,..., K$ and $\bar{q}$ represents production.

The firm’s location problem specified in equation 1 involves comparison of numerous spatial cost factors. For example, variables that affect the firm’s marginal cost function include labor market conditions, market size and accessibility, business taxes, public services, agglomeration economies, energy costs, and environmental compliance expenditures. Although these future cash outlays are obviously relevant, site-specific costs are of immediate importance to most relocating firms. This aspect may explain the recent increase in promotions and incentive packages that states have offered. Nevertheless, environmental policies may also significantly affect site-specific start-up costs. Gray (1997) discusses this idea and cites officials in the pulp and paper industry who conjecture that “the main influence of environmental regulations ... come through difficulties in getting construction permits, due either to delays in permit issuance or to uncertainty about whether the permit would be issued at all.” This phenomenon is not confined to this sector as many environmental regulations often go beyond specifying numerical discharge standards and require that a firm install specific abatement technologies prior to opening a new plant (Jaffe et al., 1995).

b. Empirical Evidence

The relationship between environmental regulation and economic activity appears to be straightforward. This intuition is supported by the theoretical model above which predicts that profit-maximizing firms will locate where the linear combination of site-specific fixed and marginal costs induced by environmental regulation are at a minimum, ceteris paribus. This notion has led researchers to pursue empirical specifications that attempt to explain changes in employment, plant expansions, relocations, or births, by using regressors that presumably affect a firm’s cost function. Empirical evidence, however, suggests stringency of environmental regulations are only weakly (or not at all) associated with decreased manufacturing activity. Although this literature has thoroughly been catalogued by Tannenwald (1997) and Levinson (1996), amongst others, a brief summary of important results is worthwhile.

Domestic evidence of the effects of environmental regulation can be split according to type of data set utilized. Duffy-Deno (1992) and Crandall (1993) use aggregate data for SMSAs (Duffy-Deno 1992) and states (Crandall 1993) and find environmental regulation does not significantly affect employment growth (Duffy-Deno 1992 and Crandall 1993) or total earnings (Duffy-Deno 1992). Other studies examine Dun and Bradstreet data. These include McConnell and Schwab’s (1990) county study of the automobile assembly industry (SIC 3711), Bartik’s (1988) state-level study of Fortune 500 companies, and Crandall’s (1993) state-wide study of plant start-ups. Although various measures of regulatory stringency are used in these studies, findings indicate that environmental regulation does not significantly alter firm location patterns. Still other studies use Census of Manufacturers data to analyze firm-level location decisions. This group includes state-level studies (Levinson 1996), and county level studies (Henderson 1996, 1997 and Becker and Henderson 1997). Although the measure of regulatory stringency again varies substantially across these studies, results from the county-level work suggest that the spatial allocation of pollution intensive
firms is affected by environmental regulation. Overall, however, results are not compelling
eough to suggest that more stringent environmental regulations systematically affects firm
location.3

A natural follow-up question in response to the general findings of the empirical
literature is “why are current studies finding that environmental regulations have negligible
effects on economic activity?” One explanation is advanced by Cropper and Oates (1992).
They argue that other firm outlays, such as expenditures for labor and energy, swamp any
effect of environmental regulations. However, using statistics from the EPA (1990) (that
suggest environmental-control costs amounted to approximately 2% of U.S. GNP in 1990 and
have subsequently increased), Levinson (1996) conjectures that these growing costs could
affect the allocation of mobile investment if spatial policies are variant enough. If spatial
heterogeneity exists in environmental policies, Levinson’s (1996b) premise would manifest
itself as inducing temporally different parameter values on the measure of regulatory
stringency in the regression equation. This is not the case, however, as comparable studies
from different decades consistently provide economically (and statistically) small effects of
environmental regulations.

This finding is potential evidence in favor of Cropper and Oates (1992) argument, but
only if environmental policies are spatially variant. Rather than presuming negligible effects
are the result of small compliance expenditures, a plausible explanation for this finding is that
spatial heterogeneity is not present in environmental policy. A potential scenario would be
that as compliance expenditures have steadily risen since the early 1970s, environmental
policies across space have become more similar, confounding any effects of increased
compliance expenditures. This premise is entirely possible given recent findings of spatial
income convergence (Gottschalk and Smeeding 1997). Indeed, assuming local income levels
partially determine environmental outcomes, income convergence may be a precursor to
conditional convergence (or convergence to a constant differential) of environmental
policies.4 Furthermore, if researchers can show that regional environmental policies have
converged through time, results would shed light on why domestic firm location studies have
consistently found environmental regulations as insignificant determinants in a firm’s
decision set, even though US firm-level compliance expenditures (for pollution abatement)
have steadily risen since the early 70s.

Complementary to this explanation is the argument that environmental compliance
expenditures are important in certain contexts. For example, it may be the case that current
empirical studies fail to pick up the effects of environmental regulation because they
misspecify the firm’s final choice set. If a firm narrows its choice set down to relatively close
geographical sites, such as between a city or one of its suburbs, rather than the 48 lower
states, state-level empirical studies may be reporting biased estimates. This argument has its
roots in the tax competition literature. A basic finding is that the smaller the area of study,
the larger the influence of tax rates. For example, intra-regional studies typically present tax
elasticities that are at least four times greater than estimated inter-regional tax elasticities
(Wasylenko, 1997). The intuition behind this phenomenon is that smaller areas tend to have
less variation in other important location factors, such as labor markets, climate, and energy
costs, which in turn accentuates any differences in tax packages.

A natural experiment to test this premise in the environmental arena is to compare
results from studies that vary the heterogeneity of the final choice set. This can be done
by comparing results from U.S. domestic studies of Standard Metropolitan Statistical Areas
(SMSAs) (Duffy-Deno, 1992), counties (Kahn, 1996; Henderson, 1996; 1997; Henderson and
Becker, 1997), and states (Levinson, 1996; Gray, 1997b) with comparable coefficient
estimates from international studies (Leonard, 1984; 1988; Tobey, 1990; Grossman and
Krueger, 1991; Low and Yeats, 1992). Evidence from the majority of studies in each
category suggests stringency of environmental regulation deters manufacturing activity, but
the magnitude of this result is generally small. Although these results indicate that
environmental regulation has little effect on manufacturing activity, if firms narrow their final choice set to smaller geographical areas than states, counties or SMSAs across the U.S., existing studies run the risk of presenting erroneous estimates. This shortcoming can be addressed in at least two ways. First, with existing county-level data, researchers could group contiguous sets of counties and analyze the location decision within each set. A second technique to further disaggregate the firm-level decision is to narrow the decision process even further, such as a choice set of a city or one of its suburbs. The analysis could then test if city ordinances regarding pollution control spur the firm to locate in the suburb. Yet, due to data availability this type of analysis may be limited to case studies.

Another possible explanation of the current empirical evidence is model misspecification. One weakness in the empirical work is the lack of controlling for increased environmental quality due to stringent environmental regulations. As previously mentioned in Section 1, although including wage rates in the specification will partially account for this problem (see, e.g. Rosen 1979; Roback, 1982, 1988; Hoehn et al. 1987; Berger and Bloomquist 1988; Gyourko and Tracy 1989a, 1989b, 1991; Voith, 1991; and Gabriel et al., 1996; amongst others), if increased environmental quality is not fully capitalized in wages, potential omitted variable bias could lead to an understatement of the effects of environmental regulation. This problem is akin to shortcomings in early studies of tax competition which failed to control for public services in the regression equation.

An innovative approach to mitigate this problem in the tax competition literature was developed by Helms (1985), who used a balanced government budget equation to allow inclusion of all revenue and expenditure items in the regression equation (also, see Evans and Karras, 1994 and Dalenberg and Partridge, 1995). Although an approach directly comparable to this technique is not readily apparent to solve the omitted variable problem in the regulatory competition literature, there should in the least be some recognition of the services that environmental factors provide. This recognition could come in many forms, including empirical techniques using instruments, simultaneous equations, or a vector autoregressive (VAR) model.

A second specification issue regards the measurement of regulatory stringency. Measures of enforcement have ranged from local spending on environmental quality control to industry-level compliance cost expenditures per dollar of value-added. Although this criticism has not gone unrecognized (see Tannenwald, 1997 and Henderson, 1996), it should be considered more closely in the final specification. Although measurement of environmental stringency is an important issue, perhaps a more glaring weakness is the failure to model the multi-dimensional nature of environmental regulation. In most cases empirical models look at environmental regulation from a single dimension and find little evidence of negative affects of environmental factors. Nevertheless, as the simple theoretical model of section 2 implies, environmental regulations can affect not only marginal costs of production but also start-up costs. Exclusion of either variable may lead to serious biases in estimated parameter values.

A final consideration regarding specification is the role that interest groups can play in the siting process. Hamilton (1993) hypothesizes that pure discrimination and variations in the propensity of communities to engage in collective action to oppose the location of potential polluters effects the distribution of polluters. The first theory is commonly termed environmental racism and has gained enough acceptance to merit creation of an Office of Environmental Equity, which is overseen by the EPA. Also, Hamilton (1993) and Arora and Cason (1997) have recently presented results that suggest collective action significantly effects communities’ exposure to pollution. Hence, empirical specifications that neglect these factors run the risk of biased parameter estimates.

A few other considerations that potentially explain the lack of robust evidence regarding the effects of environmental regulation on economic activity are: (1) existing
market structures—many regulated industries are natural oligopolies which are much less affected by environmental regulation since the regulatory tax incidence may ultimately fall on consumers; (2) endogeneity issues—studies failing to lag independent variables, which may lead to biased estimates as a large influx of firms may affect the regulatory regime; and (3) geographically immobile firms—incidence of environmental regulation may fall on firms which are not footloose enough to migrate and take advantage of lax regulation in other locations. These (and other) arguments are summarized in Levinson (1996) and/or Tannenwald (1997).

c. Where Do We Go From Here?

Besides issues discussed above, many potential research topics remain in the interjurisdictional competition area. For example, Deily and Gray (1991), Crandall (1993), and Kahn (1996) have recently presented results that support an interesting premise relating local business cycles to local environmental regulations. For example, their findings suggest that oftentimes regulators advance special deals which stretch out compliance schedules of individual industries, particularly those industries which are large employers. Hence, enforcement may not be as random as first perceived, as regulators’ policing of polluters appears to recognize the trade-off between the environment and jobs. On the other hand, findings from these papers suggest that employment growth in polluting industries is inversely related to regulatory stringency. Together, these findings suggest that environmental regulation may tend to smooth out the local business cycle—periods of growth are slowed by environmental regulations while recessionary periods are made less severe because environmental regulations tend to protect the interest of large, existing firms (see, also, Pashigian, 1985). Nevertheless, this important phenomenon deserves more attention.

This research agenda could commence with determining firm hazard rates across counties of attainment/nonattainment. These results could then be compared to the intuition of Stigler (1971), who argued that a plant with an older capital stock is more likely to be closed than a plant with a newer capital stock, ceteris paribus, simply because major re-investment decisions should arise in the older plant first. However, there is a caveat to Stigler’s argument—we are not in a world of ceteris paribus. For example, current environmental regulations are written in a manner that regulates new sources of pollution much more stringently (new source bias) than older sources (older firms have been “grandfathered” into less restrictive standards and have a relative advantage).

Other related studies also abound. In a world of environmental regulation, does size of a plant affect firm hazard rates; location; or expansion? Also, do environmental regulations have heterogeneous effects across industry? If environmental regulations do affect any type of economic activity, the relationship should be uncovered in studies of pollution-intensive industry, such as chemicals, primary metals, or papers.
References


Pashigian, Peter, “Environmental Regulations: Whose Self-Interests are Being Protected?” Economic Inquiry 23 (October 1985), 551-84.


Endnotes


2. Another technique to measure the effects of environmental regulation on economic activity is surveys. However, given the lack of incentive compatibility that plagues hypothetical surveys, we suppress discussion of the survey literature. Nevertheless, Levinson (1996a) provides a good review of current survey evidence.

3. Note that general findings from international location studies imply similar results (see, e.g., Leonard, 1984; 1988; Tobey, 1990; Grossman and Krueger, 1991; Low and Yeats, 1992a; 1992b).

4. Convergence up to a constant differential, or conditional convergence, is stressed since optimal environmental regulation should take into account other factors that affect potential damages. These factors include assimilative capacities, demographic characteristics, geographic characteristics, etc.

5. This is similar to McFadden’s conditional logit model, except here the final choice set is not 9 randomly selected counties, but n contiguous counties.

6. This omitted variable bias is attenuated in panel data contexts in which quality of the environment is most likely captured by the fixed effects.

7. Wages and housing costs probably do not capitalize all the effects of environmental quality. This is illustrated by Schmidt (1958), who noted that many of Pittsburgh’s largest corporate employers were planning on moving in 1945 because their wives “didn’t want to live and raise their families under...prevailing environmental conditions.”

8. This concern may be mitigated, however, as most specifications include per capita income levels, which may be a proxy for community groups and racial composition.

9. Work in this area has been initiated by Becker and Henderson (1997). However, an important caveat is mentioned by Levinson (1996), who notes that many times plants will employ skeletal work crews to avoid regulatory costs of shutting down. It appears that Becker and Henderson’s (1997) data set does not overcome this shortcoming.

10. A few recent studies that analyze environmental regulations effects across heterogeneous industry include Levinson (1996b) and Gray (1997b).