AgEcon SEARCH
RESEARCH IN AGRICULTURAL & APPLIED ECONOMICS

The World's Largest Open Access Agricultural & Applied Economics Digital Library

This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.

Help ensure our sustainability.

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

Using DEA and VEA to Evaluate Quality of Life in the Mid-Atlantic States

Asst. Prof. Elizabeth Marshall*
Pennsylvania State University
Department of Ag. Econ. and Rural Soc.
Armsby 112F
University Park, PA 16802
LizMarshall@psu.edu
Ph: 814-863-8634
F: 814-865-3746

Prof. James Shortle
Pennsylvania State University
Department of Ag. Econ. and Rural Soc.
Armsby 112B
Pennsylvania State University
University Park, PA 16802

Paper presented at the 2004 American Agricultural Economics Association Meeting, Denver, CO.

* Corresponding author

The authors wish to acknowledge the research assistance of Yun Cai at Penn State University.

Using DEA and VEA to Evaluate Quality of Life in the Mid-Atlantic States

Abstract: In this study we use data envelopment analysis (DEA), and an extension of DEA called value efficiency analysis (VEA), to explore the "production" of quality of life within counties in the mid-Atlantic region and the extent to which production frontiers and efficiency differ between rural and urban counties. These methods allow us to identify counties that are inefficient in their quality of life production, and to rank those counties according to their distance from a performance standard established by other observed counties (using DEA), or by a single unit designated as "most preferred" (using VEA) .

Key Words: Data envelopment analysis, value efficiency analysis, quality of life.

Introduction

Speculation regarding the relationship between the attributes of a community and the quality of life experienced by its residents has gained vigor in recent years. In contemporary discussion, the issue arises most visibly in debate about the dispersed, auto-dependent development patterns known as urban sprawl, and in burgeoning interest in such concepts as smart growth and liveable or sustainable communities. Additionally, observed changes in migration patterns nationwide, together with increasing interest in the role that natural amenities play in residential choice behavior, highlights the importance of understanding how residents value the various attributes associated with different types of communities. Given the abstract, multi-dimensional nature of the underlying concept of quality of life, however, quantifying relationships between community attributes and quality of life poses significant challenges.

Efforts to measure quality of life confront two equally challenging tasks. One is to identify a set of indicators that represent appropriate dimensions for measuring quality of life. Such indicators are often selected to reflect economic, social, and environmental factors. The second task is to aggregate such indicators into a composite measure that can be used to differentiate communities along a quality of life spectrum. One of the most commonly used methods for evaluating quality of life has been the hedonic price method. This method is based on theoretical work by Rosen (1979) and Roback (1982) suggesting that, at a labor- and land-market equilibrium, the value of regional amenity and quality of life factors should be capitalized into regional wages and rents (Deller, 2001). Differentials among regional wages and rents should therefore reflect differences in quality of life, and these differentials can be used to estimate the values attached to each amenity factor. Blomquist, Berger and Hoehn (1988) used hedonic wage and rent models to estimate implicit amenity prices for a variety of regional climatic, environmental and urban factors; these prices then served as weights in a quality of life index applied to urban counties. A more recent application of this technique is Gabriel et al. (2003), whoe extend the hedonic equation system to include the price of locally traded goods

other than housing, and apply the analysis to pooled cross section and time series data on a large variety of amenity and quality of life variables. This allows them to not only estimate the implicit price of amenity factors but also to construct a state-level quality of life index based on these amenity weights.

Hedonic amenity weight estimates, and the quality of life indices that arise from them, are sensitive to the specification of the functional form linking amenities with existing wage and income differentials. Non-parametric quality of life indices avoid this issue. In an effort to eliminate completely the need for "ad hoc" selection of quality of life indicators, Douglas et al. (1993) appeal to the legacy of Tiebout (1956) in arguing that migration patterns should reflect quality of life differentials; mobile residents will "vote with their feet" for those communities with high quality of life. They therefore construct a quality of life index based on the presumption that the probability of a resident moving from location A to B will depend on the degree to which the quality of life in location B exceeds that of location A (Douglas, 1993). This index allows them to rank locations in a relatively objective manner, but they do not attempt any further examination of the specific factors contributing to these quality of life rankings.

In this study we use data envelopment analysis (DEA), and a recent extension of DEA called value efficiency analysis (VEA), to explore the "production" of quality of life within counties in the mid-Atlantic region of the United States using counties as the units of analysis. These methods allow a non-parametric approach to ranking communities and analyzing the contributions of different factors to quality of life production. DEA is a non-parametric frontier analysis method that was originally developed to analyze the performance of organizations whose goals are not limited to profit maximization (Charnes et al., 1978). The methodology has been applied to issues as diverse as fisheries management, health care provision, national defense, and the socio-economic performance of nations (Bowlin, 1998; Golany et al., 1997; Walden & Kirkley, 2000).

Using linear programming, DEA creates a frontier of efficient units that envelopes other,

relatively inefficient units. Various measures of inefficiency are available, based on different methods for measuring the distance from a unit's observed production point to the efficient frontier. DEA is flexible in that it does not require specification of an underlying production relationship between inputs and output, it is able to incorporate inputs and outputs that are measured in different units and at different scales, and it is able to accommodate multiple inputs and multiple outputs with minimal value judgments placed on the relative "worth" or "cost" of these inputs and outputs. The first and second features make DEA appealing for evaluating the performance of communities in providing quality of life using varying measures of economic, social, and environmental inputs and outputs. In using DEA for that purpose, the third feature represents an innovative approach to the aggregation of quality of life indices. Essentially, it provides an objective procedure for weighting inputs and outputs that requires the analyst to assume no more than that outputs should be maximized and inputs should be minimized. Since value judgements about the relative worth of alternative community attributes create potential mine fields in quality of life assessment, this third feature is one that offers some appeal for objective assessments.

Extensions of the DEA technique that retain the first two features noted above but that allow for stronger value judgements to be imposed regarding the relative desirability of inputs and outputs have been developed. Value efficiency analysis (VEA) is one such extension that is appealing in that it does not require those evaluating quality of life to explicitly assign weights, or relative weights, to inputs and outputs. Instead, evaluators need only to select a "reference community" against which other communities will be measured. The levels of inputs and outputs of that reference community establish implicit constraints on the weights that can be assigned to inputs and outputs in the remaining communities.

In the following section we briefly explain the basic structure of DEA and VEA. We then describe the quality of life model, explaining the relevant dimensions and the data used in our efficiency analysis, and present results. In doing so, we will explore the hypothesis that a

fundamentally different efficiency frontier exists for rural counties than for urban counties; if this is the case, a county's performance should be measured relative to only those counties that share its rural/urban classification.

Data Envelopment Analysis and Value Efficiency Analysis

We envision counties as entities that make a collection of development decisions that in aggregate produce quality of life for its residents. Desirable outcomes that are created, such as employment opportunities and high quality educational systems, may be accompanied by undesirable outputs such as crime and pollution. In choosing a development path to maximize quality of life for its residents, counties would like to maximize the desirable outputs and minimize the undesirable outputs. Counties that are relatively efficient at producing high quality of life will produce relatively more desirable outputs per unit of undesirable output than counties that are relatively inefficient.

Data Envelopment Analysis

DEA provides a uniquely flexible way to model the scenario described above. Fundamental to our measurement of efficiency is the proposition that different development paths lead to communities with different combinations of attributes, and that overlying all of these possible combinations is a community attribute frontier that represents maximum achievable performance along the many dimensions making up quality of life. Efficient counties lie along the frontier and produce a higher ration of desirable outputs to undesirable outputs than the inefficient counties that lie within the frontier. Using these frontiers as a standard for judging the relative performance of counties in producing quality of life requires few value judgments about the relative worth of various desirable or undesirable community characteristics. DEA makes only the weak but reasonable assumption that communities prefer to have more of "good" development outcomes (e.g. natural amenities, literacy, affordable living) and fewer of "bad"

development outcomes (e.g. pollution and poverty).

Using terms borrowed from more traditional production relationships, in our analysis entities to be minimized (undesirable outcomes) will be referred to as inputs, and entities to be maximized (desirable outcomes) will be referred to as outputs. Of course, no direct production relationship exists among these factors– inputs are not literally consumed to produced outputs– but DEA is a non-parametric methodology and requires no assumptions about the form of the underlying relationship connecting inputs and outputs.

The most basic DEA formulation evaluates the relative efficiency of a production unit by estimating for each unit a measure of $\frac{weighted\ outputs}{weighted\ inputs}$ (Cooper et al., 2000). A unit with an efficiency ratio of 1 is efficient, while one that uses relatively many inputs, or produces relatively few outputs, will be found inefficient with a ratio of less than 1. One fundamental innovation that DEA offers is an objective method of determining what weights will be assigned to outputs and inputs in determining this ratio. The procedure computes set of weights for each decision-making unit (DMU) that maximizes its efficiency ratio, subject to the constraint that the efficiency ratio calculated at that set of weights does not exceed one *for any DMU in the data set*. Accordingly, the assigned weights vary by DMU and are derived from the data such that each DMU is allowed to be as efficient as possible relative to the other DMUs.

The original formulation for DEA can therefore be expressed as the following fractional programming problem, which must be solved for all n DMUs in the data set:

$$\max \quad \theta = \frac{u_1 y_{1o} + u_2 y_{2o} + \ldots + u_s y_{so}}{v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo}} \qquad [1]$$

$$\text{subject to} \quad \frac{u_1 y_{1j} + u_2 y_{2j} + \ldots + u_s y_{sj}}{v_1 x_{1j} + v_2 x_2 + \ldots + v_m x_{mj}} \leq 1 \qquad j = 1, \ldots, n$$

$$v_i \geq 0 \qquad i = 1,\ldots,m$$

$$u_k \geq 0 \qquad k = 1,\ldots,s$$

where y_{ik} denotes the level of output k for DMU j, x_{ij} denotes the level of input i for DMU j, u_k represents the weight assigned to output k for the base decision-making unit, v_i represents the weight assigned to input i for the base decision-making unit, and θ is the efficiency measure. Note that the base decision-making unit for each linear programming problem is denoted with the subscript o. The above problem can be easily reformulated as the equivalent linear programming problem below:

$$\text{max} \quad \eta = \mu_1 y_{1o} + \mu_2 y_{2o} + \ldots + \mu_s y_{so} \qquad [2]$$

$$\text{subject to} \quad v_1 x_{1o} + v_2 x_{2o} + \ldots + v_m x_{mo} = 1$$

$$\mu_1 y_{1j} + \mu_2 y_{2j} + \ldots + \mu_s y_{sj} \leq v_1 x_{1j} + v_2 x_{2j} + \ldots + v_m x_{mj}$$

$$j = 1,\ldots,n$$

$$v_i \geq 0 \qquad i = 1,\ldots,m$$

$$\mu_k \geq 0 \qquad k = 1,\ldots,s$$

DMUs with weights that yield an efficiency rating of 1 define an efficient production frontier. Those with efficiency ratios less than 1 fall some distance from the frontier. A measure of this distance is a measure of the inefficiency of the DMU in question. For this formulation, the measure of distance is a radial measure indicating how the output/input ratio would have to change in order for that DMU to be considered efficient (i.e. to occupy a position on the efficient frontier).

To illustrate, suppose we have five DMUs producing two outputs using one input. The production possibilities frontier, assuming a constant input level equal to 1, is shown in figure 1. DMUs one through four lie on the frontier, and are therefore efficient (i.e. distance to the frontier

is zero). The solution for the above LP model would therefore yield and efficiency ratio of one for each of these units. DMU 5, however, lies within the frontier. The solution of the LP problem for DMU 5 yields an efficiency ratio of η=.527. In figure 1, this result is the ratio of distances $\frac{OB}{OA}$, where O refers to the origin, and points A and B are as indicated on the graph. The ratio indicates that the output/input ratio for DMU 5 is 52.7% of what it is required for technical efficiency. Accordingly, to become efficient, DMU 5 can either reduce inputs to 52.7% of their current levels, or increase the level of each output produced to $\frac{y_{5i}}{.527}$ (i=1,..., s). The latter adjustment, for instance, would move DMU 5 to the point marked A (1.89, 5.68) on the efficiency frontier.

To assure validity of the results, it is important to measure the performance of DMUs relative to an appropriate frontier. With our quality of life analysis, for instance, it may not be appropriate to measure the distance of inefficient rural counties from a frontier partially formed by efficient urban counties since the production relationship for urban and rural counties and their efficient frontiers may be fundamentally different. To illustrate, consider figure 2, which depicts three frontiers for two abstract outputs– environmental quality of life and economic quality of life. The frontier for urban communities only implies that urban regions are capable of producing very high economic quality output, but that their environmental quality output is limited. The frontier for rural communities implies the opposite; rural regions are able to produce high environmental quality, but it is not technically possible (according to observed data) to also produce high economic quality of life. The remaining frontier represents an overall frontier that would be computed by combining rural and urban communities.

To illustrate the problems inherent in using inappropriate frontiers for efficiency measurement, consider point A, which represents an urban community. This community will

appear considerably more efficient as a production unit when measured relative to other urban regions (as shown along the line from the origin to the urban frontier) than when measured against rural regions or against a composite frontier. If it is the case that urban regions are fundamentally incomparable to rural regions in the provision of environmental amenities, then the two groups should be used to generate different efficiency frontiers. In this study, we use our efficiency results to explore whether there is evidence that such differential frontiers exist.

Value Efficiency Analysis

DEA evaluates DMUs with respect to the relatively weak criterion of technical efficiency. However, it can be useful to evaluate performance under stronger assumptions about the relative value of inputs and outputs by imposing logical restrictions on the weights allowed. These restrictions reflect judgements that exist independent of the data and that impose restrictions on efficiency measurements based on pre-existing perspectives about the relative importance of individual inputs and outputs or prior views on what constitutes an efficient or an inefficient DMU (Allen et al., 1997). We can then ask, how do communities perform contingent on a set of judgements about the relative value of different community outcomes?

Several methods of imposing judgements via weight restrictions have been developed in the literature, but in this study we will implement value efficiency analysis– a recent extension of DEA that allows the decision-maker (DM) to incorporate into the efficiency analysis some information about preferences among production outcomes. The task of assigning constraints to weights or relative weights in DEA can be quite difficult and somewhat arbitrary, and is complicated in cases where the inputs and outputs are measured in different units. Rather than requiring the decision-maker to set such constraints, VEA simply requires the decision-maker to designate one DMU as a "most preferred solution" (MPS), denoted u*, which will be used as the standard against which other DMUs are measured. The decision-maker is assumed to have a value function that is pseudoconcave, strictly increasing in outputs and strictly decreasing in

inputs, and that reaches a local maximum at the point u* on the efficient frontier (Halme et al., 1999; Korhonen et al., 2001). That point u* is therefore assumed to lie along a value contour that is unknown but, given the above assumptions, can be linearly approximated by the hyperplane tangent to the efficient production surface at u* (Halme et al., 1999). Measurements of value efficiency evaluate how far a DMU lies from this approximated value contour, rather than from the technical efficiency frontier itself.

VEA is illustrated in figure 3. Assume that from among the DMUs previously identified as efficient (DMUs 1 through 4), DMU 3 has been identified as the most preferred solution. It is assumed that there is an underlying value function whose exact functional form is unknown; a single isovalue curve is illustrated on the figure passing through the most preferred point. A value efficient measurement with full information would indicate how far each unit lies from the isovalue contour itself. However, in VEA the exact functional form of the underlying value function is unknown, as is the exact location of the isovalue contour. An approximation is used instead. Given the above assumptions about the utility function, a plane formed parallel to an adjacent facet on the efficient frontier will encompass that isovalue curve, and is therefore an upper bound for the value efficiency measurement. The dotted line indicated in figure 3 represents such an approximation, and a lower bound for the inefficiency measurement can be found using the distance of value-inefficient points to this line. Note that DMU 1, which was efficient according to DEA, would be found to be value-inefficient using VEA because its mix of outputs, though technically efficient, is dissimilar to that of the most preferred solution– DMU 3. DMU 5, which was found to be DEA-inefficient, would be assigned an even smaller efficiency measurement using VEA because efficiency is now measured relative to the approximated isovalue curve, rather than the technical efficiency frontier. In this study, we will expand upon the technical efficiency analysis provided by DEA to explore the value-efficiency of DMUs relative to a “most preferred solution” in the provision of quality of life, as well as the sensitivity of those value-efficiency scores to selection of the MPS.

Model and Data

In this section we describe the quality of life indicators used in our analysis and describe in detail the procedures used to estimate the efficiency of quality of life provision by counties in the mid-Atlantic states.

Quality of Life Indicators

Quality of life indicators are broadly categorized into three dimensions– social, environmental, and economic– for which we wanted representation among our quality of life variables. There are numerous possible variables within each of these categories. The final selection of the model variables was based on several considerations and limitations: an extensive review of the quality of life literature, the availability of complete data at the county level, a desire to balance the representation of the three quality of life dimensions in the final model, and a tradeoff that the DEA methodology generates between the number of inputs and outputs that can be included and the discriminatory power of the model. As the number of inputs and outputs included in the model increases, efficiency scores of all DMUs in the data set increase. It is important, therefore, to ensure that relevant variables be included in the model but to take steps to eliminate variables that contain duplicate or unrelated information.

The variables included in the model are shown in Table 1. Inputs to the model are considered to be those development outputs that counties would like to minimize– including factors that might be characterized as negative "side effects" of development, such as cancer risk, and factors that describe positive development outcomes, but for which smaller means better, such as teacher/pupil ratios. Model outputs are those outcomes that counties would like to maximize, such as affordability or percentage of the population with a high school degree. Note that the latter category could also include factors that are not literally "produced" in the development process, such as acres of natural areas– because no underlying production function is assumed, the model does not distinguish between desirable characteristics that are actually

produced, and those that merely survive the production process unscathed.

In many production scenarios, variables exist that affect production but are nevertheless out of the control of the DMU, for example weather in agricultural production models. Such variables are called "non-discretionary" or "fixed" variables in the DEA literature. We include a single non-discretionary variable– the amenity index produced by the Economic Research Service to describe a region's climate, topography, and water area. This variable has been found to be significant in explaining rural migration patterns over the past couple of decades (ERS, 1999). The amenity index is assumed to influence quality of life, and therefore to influence how counties rank with respect to efficiency, but it cannot be manipulated in order to improve an efficiency rating. A county with high amenity value, for instance, may score quite high with respect to quality of life, but an inefficient county with low amenity value cannot change that value to increase efficiency; the remaining output variables may have to be improved by an even greater amount to compensate for the fixed (and poor) amenity value.

The cancer risk index provides for each county a measure of the risk of cancer-related health effects resulting from exposure to air toxics. The data set is based on 1996 emissions data collected through the EPA's National-Scale Air Toxics Assessment. The percent of land area developed is calculated from the 1992 National Land Cover Data. The data on teacher/pupil ratios for the 2000-2001 school year is obtained from the National Center for Education Statistics. The percent of residents below the poverty level and percent of residents 25 years and older who hold a high school degree or equivalent are all obtained from 2000 census data. The number of recreation and entertainment establishments per developed square mile is calculated from 2001 County Business Pattern Data (NAIC code 71) combined with data on developed area from the National Land Cover Data.

The affordability index is a variable constructed with information about both cost of living and median household income. The Family Economic Self-Sufficiency Project has generated for most states figures indicating the hourly or annual wage that would be necessary

for families of various sizes to achieve self-sufficiency. These figures, called the "self-sufficiency standard", incorporate information regarding the costs of housing, food, childcare, transportation, healthcare, taxes, and miscellaneous other essentials such as clothing, household items, personal hygiene, etc. The standards presented for each family size vary by county, city, or metropolitan area within a state.

We selected as a representative household a two-parent household with one pre-school child and one school-age child. In those cases in which the self-sufficiency standard for a city or other high-cost area was calculated separately from the surrounding county, we created for the county an aggregate self-sufficiency standard by taking a weighted average of the standards for the county and the city or cities it encompasses, using total number of households in each geographical area as the weight. The result was a county-level index indicating what the annual household wage would have to be for a representative household (as described above) to be self-sufficient.

The affordability index was calculated as the ratio of median household income (from Census 2000) to the self-sufficiency index described above. A ratio of greater than 1 indicates that more than half of the households in a county have adequate income to be self-sufficient. Such a county is considered more affordable than one in which the affordability index is less than one, suggesting than more than half of the households do not earn enough income to be considered self-sufficient. The affordability index therefore captures two different factors affecting quality of life– the expenditures necessary to live in that county, and the availability of economic opportunities to provide income to cover the cost of those expenditures. In theory, a county could increase its affordability index in one of two ways– by lowering costs or by increasing incomes.

For all of the variables described above, data from Virginia's independent cities were combined with surrounding counties according to the guidelines presented in the 1989 ERS County Typology Codes.

DEA and VEA methods

The prior section provided a brief overview of the most fundamental DEA model. We use several extensions to this original model. The model described above, for instance, assumes an underlying constant returns to scale in the relationship between inputs and outputs. We use a less restrictive model that allows for the existence of variable returns to scale in the production relationship– the BCC model developed by Banker et al. (1984). The original model also assumes that all inputs and outputs can be controlled by the decision-maker and are therefore able to be adjusted in arriving at an efficient production point. To accommodate a non-discretionary output, we used a further extension as described in Banker and Morey (1986). The model that we use is therefore an output-oriented, customized BCC model, as shown below[1]:

$$\text{Minimize} \quad \sum_{i=1}^{M} v_i x_{i0} - \sum_{r \in y_F} \mu_r y_{r0} + u_0 \qquad [3]$$

$$\text{subject to} \quad \begin{aligned} & \sum_{r \in y_D} \mu_r y_{r0} = 1 \\ & \sum_{r \in y_D} \mu_r y_{rj} - \sum_{i=1}^{M} v_i x_{ij} + \sum_{r \in y_F} \mu_r y_{rj} - u_0 \le 0 \qquad \forall j \\ & v_i \ge 0 \qquad \forall j \\ & \mu_j \ge 0 \qquad \forall j \end{aligned}$$

where F denotes the non-discretionary output. The series of linear programming problems

[1]Note that the formulation originally discussed was an input-oriented formulation, which selects weights to maximize output subject to a fixed input level. The output-oriented formulation shown here is an output-oriented formulation that selects weights to minimize inputs subject to a fixed discretionary output level.

described above was programmed in Excel. The linear programming problems described above could, like other linear programming problems, alternatively be formulated and computed using the dual formulation. Because this alternative formulation is more commonly solved in the literature, that formulation has become accepted as the primary DEA formulation, and the formulation described in equations section [3] has become accepted as the dual. The primary formulation is generally considered easier to solve computationally because it has far fewer constraints; the dual described above has at least as many constraints as DMUs. However, the primary formulation has far more decision variables (again, at least as many as there are DMUs), and in this application the number of decision variables in the primary problem exceeded Excel's limitations on allowable problem size.

The value efficiency analysis involves only a small change in model formulation and is computationally comparable to the DEA. The dual VEA formulation that we used is shown below:

Minimize $$\sum_{i=1}^{M} v_i x_{i0} - \sum_{r \in \gamma_F} \mu_r y_{r0} + u_0 \qquad [4]$$

Subject to:

$$\sum_{r \in \gamma_D} \mu_r y_{r0} = 1$$

$$\sum_{r \in \gamma_D} \mu_r y_{rj} - \sum_{i=1}^{M} v_i x_{ij} + \sum_{r \in \gamma_F} \mu_r y_{rj} - u_0 \le 0 \qquad \forall j$$

$$\sum_{r \in \gamma_D} \mu_r y_{rk} - \sum_{i=1}^{M} v_i x_{ik} + \sum_{r \in \gamma_F} \mu_r y_{rk} - u_0 = 0$$

$$v_i \ge 0 \qquad \forall j$$

$$\mu_j \ge 0 \qquad \forall j$$

The computational modification of DEA to form VEA simply involves an additional constraint ensuring that the reference unit, indexed above by k, remains efficient at all sets of weights used by other counties in calculating their value efficiency measurements. In other words, under VEA, no county may use a set of weights in determining their efficiency score that would render the

reference unit itself inefficient. Such a restriction can considerably restrict the weights that non-reference counties can adopt.

DEA Results:

To determine whether to combine or separate urban and rural counties when measuring relative efficiency, we first test whether the frontiers for these groups differ in a significant and systematic way. Counties are classified as urban or rural according to their Beale codes. Counties with a Beale code of 0-3 are considered urban, while those with a Beale code of 4-9 are considered rural.

To test for differences in the efficiency frontiers, we first use DEA to construct efficient frontiers for each group. Next the results of the DEA analysis are used to project those units that are inefficient onto the efficient frontier, creating new sets of data that consist of both currently efficient units and formerly inefficient units that have been adjusted to efficiency. When run only within its rural or urban category, every unit in the new data sets would now be rated efficient. We then pool the new data sets back into one large data set and perform a combined DEA. The results of the DEA on the combined data are then ranked according to their efficiency ratings. If there are no significant differences in the two frontiers, there should be no significant difference in how urban and rural counties are ranked in the combined model. As suggested in Cooper et al. (2000), we then use the Wilcoxon-Mann-Whitney test to test for significant differences in the rankings of rural and urban counties.

The T statistic for that test is:

$$T = \frac{S - \frac{m(m+n+1)}{2}}{\sqrt{\frac{mn(m+n+1)}{12}}} \qquad [5]$$

where S denotes the sum of the rankings of the urban units, m denotes the number of urban units,

and n denotes the number of rural units. The T value that we derive in testing the hypothesis of no significant difference between urban and rural rankings is T=-10.64. For a two-tailed test the critical level associated with α=.05 is $T_{.025}$=1.96. We therefore reject the hypothesis of no significant difference in efficiency frontiers. Because there is evidence that the production frontier differs between rural and urban counties, the appropriate efficiency measures are those that we originally estimated in calculating the DEA model separately for the two groups.

The implication of the difference in frontiers is that rural and urban counties use significantly different production technologies in generating quality of life for their residents. In general, this may be because the frontiers intersect (as shown in the graph above), or because one of the frontiers is completely enclosed within the other. If the latter scenario were the case, however, one would expect to find that, when the model is run on all of the data together, all of the members of one group are found to be inefficient relative to the other group. This is not what we observe. When the full model is run on the projected data (as described above), the resulting frontier is predominantly urban, but also contains a small number of rural counties; this suggests that although most of the rural frontier is enclosed within the urban frontier, the frontiers intersect along one or more dimensions. Nevertheless, in general, even when all units are operating efficiently according to their respective technologies, urban counties will tend to dominate over rural counties in the provision of quality of life.

Given the adjustments for return to scale, and the multiple dimensions involved, it is not possible to illustrate the actual frontiers. However, aggregate data differences between the two groups may provide additional clues about their relative "productivity". Average input and output values for the two groups are shown in table 2. Not surprisingly, rural counties appear to outperform urban counties with respect to the environmental quality inputs, while urban counties outperform rural counties along the socioeconomic dimensions of affordability, poverty, percentage with a high school degree, and access to arts, recreation, and entertainment facilities. Although these averages are useful for highlighting the relative strengths of the different groups,

it is important to recall that the frontiers themselves are not generated by average performance– the frontier is composed of reference extremes from either group.

It is also interesting to explore the averages of the projected values that are used to generate the rural/urban frontiers for testing– i.e. average performance within a group assuming that all units are performing efficiently. These averages are shown in table 2 as well. Note that units are projected onto an efficiency frontier taking inputs, and non-discretionary outputs, as fixed, so that only the values of the discretionary outputs change. Although the frontier reduces the gap in performance that exists along some of the dimensions, for the most part average relative differences between the groups have been maintained. Together with the frontier test above, this finding implies that in calculating their efficiency measures, rural counties tend to be measured relative to a higher standard than urban counties along the environmental dimensions, but relative to a lower standard along the socioeconomic dimensions.

The actual efficiency values calculated for the counties in the mid-Atlantic region are shown in figure 5. Recall that these are relative efficiency measures, and, as explained above, they are not measured relative to all other counties but only to other urban or rural counties, as appropriate. This figure suggests that the counties with highest efficiency are scattered throughout the mid-Atlantic region, with the areas of lowest technical efficiency concentrated in West Virginia and Virginia. These counties tend to perform poorly along a number of indicators– poverty level, percent high school graduates, and affordability, in particular.

VEA Results

In measuring technical efficiency via DEA, we have made no value judgements regarding the importance that should be given to the various dimensions of quality of life in determining a county's performance; in determining their most advantageous set of weights counties may specialize in producing certain outputs, or in economizing on certain inputs, and each input and output is given equal importance. A county may, for instance, be considered efficient because it

has a high affordability index– this county specializes in producing income opportunities or in keeping the cost of living low. Another county may earn an efficiency score of one because it has an extremely low pupil/teacher ratio– this county specializes in providing a high-quality education. Both counties earn the highest efficiency rating possible, and no further distinction between them is made based on the original DEA model.

Although people likely consider several, or all, of these "products" when evaluating a county based on quality of life considerations, is it reasonable to assume that they are all equally important, and interchangeable, in that decision? If not, weight restrictions of some kind must be added to the model to impose additional structure on the performance measure. In this study, we incorporate additional information on preferences for different quality of life dimensions through the use of value efficiency analysis, as described above. The literature contains some references addressing the selection of the "most preferred solution" (MPS) through analysis of the data itself (Korhonen et al., 1998). Our preference, however, was to use external data to provide additional information about quality of life comparisons among the candidate counties.

By assumption, the most preferred solution must lie on the production frontier, and therefore are efficient according to the original DEA measurements. We therefore restrict our consideration of possible MPSs to those counties that belong to the original DEA-efficient sets in their respective groups. We then used aggregate net migration rates to derive conclusions about migrants' judgments regarding relative level of quality of life among those efficient counties. We combined 1995-2000 inter-county migration data from the census with census estimates of the 1995 population to derive net migration rates for all of the urban and rural efficient counties. The county in each set with the highest net migration rate was selected as that data set's most preferred solution.

For the rural data, Worcester County, Maryland, emerged as the county with the highest net migration rate. Worcester County is the eastern-most county in Maryland; it is located on the Atlantic Ocean and runs from the Delaware border on the north to the Virginia border on the

south. Ocean City, Maryland, is part of Worcester County, as is the Assateague Island National Seashore, and much of Chincoteague Bay. Its land area is 456 square miles, almost 3% of which is in urban or suburban land use, and according to Census 2000 is home to 46,543 people. Despite Worcester County's coastal location, it's amenity index is only slightly above average. Relative to other rural counties, however, the county scores quite high for concentration of arts and entertainment establishments.

Among the urban efficient counties, Loudoun County, Virginia emerged as the county with the highest net migration rate. Loudoun County is located in the Washington DC metropolitan area. Although Loudoun County is currently relatively unpopulated– about 220,000 residents share 517 square miles– it has been growing rapidly, particularly in the eastern portion of the county. Fairfax County, which lies to the east between Loudoun County and Washington, DC, is the most populated county in Virginia, with slightly more than a million people squeezed onto 395 square miles. Loudoun County outperforms the average urban county along every dimension except the amenity index; it performs particularly well with respect to percent of residents below the poverty level and percent of residents with a high school degree or equivalent.

Relative to the DEA scores, VEA scores can only stay the same or drop; it is not possible to be rated as more efficient when a reference community is used than when the frontier itself is used as the performance standard.. The value efficiency measures for the mid-Atlantic region are calculated using Worcester County as the MPS among rural counties and Loudoun as the MPS among urban. Figure 6 illustrates the extent of the drop in efficiency ratings when VEA rather than DEA is used. The efficiency scores of rural counties are only slightly influenced by the use of Worcester County as a reference county, but the efficiency scores of certain urban counties drop significantly when held against Loudoun as a reference.

In the VEA analysis, counties are restricted to a set of input/output weights which, when applied to the reference county, would allow the reference county to maintain an efficiency score

at, but not above, a level of 1. In general, therefore, the counties with the highest value efficiency scores are those whose combinations of inputs and outputs are most similar to those of the counties identified as the most preferred among urban and rural counties. The response of the VEA analysis to selection of an MPS therefore depends on the extent to which all the other counties are able to perform well under various combination that are favorable to the reference county.

To illustrate a scenario where that is not the case, consider Amelia County, Virginia. Amelia County is located only 35 miles from Richmond, VA, and its proximity to that urban center earns it a Beale code that classifies it as urban. However, relative to other urban counties, Amelia is unpopulated and undeveloped. In fact, its mix of inputs and outputs is more comparable to the rural counties than to the urban ones. It outperforms the urban average on only the four input dimensions; its performance along the four output dimensions is not impressive. In achieving its DEA-efficient rating, Amelia very heavily weights both the poverty level input and the % developed land area input. It is unable to retain that rating under the VEA analysis, however, as the comparison to Loudoun County requires it to lower the weight placed on those inputs, and, because its performance along the output dimensions is poor, it is unable to compensate by more heavily weighting its output dimensions as Loudoun County does.

The VEA-efficiency landscape would look much different if Amelia County were used as the urban reference county rather than Loudoun County. Amelia County comes in second among the urban efficient counties in terms of net migration rate, so it might have been considered a reasonable candidate for the reference community. In order to illustrate the sensitivity of the VEA results to the choice of the most preferred solution, consider what the efficiency loss figures would look like if the counties that came in as runner-ups in the net migration criteria were used rather than the winners. Figure 7 illustrates the efficiency figures if Amelia County, Virginia, is used as the urban reference county and Floyd County, Virginia, is the rural reference county. These are not efficiency loss figures, they are the actual efficiency scores; clearly the efficiency of

the urban counties plummet, many of them to close to 0. Floyd County has a less dramatic, but still noticeable, effect on the efficiency scores of rural counties.

As figure 7 indicates the VEA results are highly sensitive to choice of the most preferred solution within each group. One way to anticipate the response of the VEA analysis to a particular MPS is to characterize the candidate MPS counties on a continuum from "generalist" to "niche" performance. Generalist counties are counties that earned their DEA-efficient rating by performing above average on a large number of dimensions, while niche counties are able to earn a DEA-efficient rating by specializing their performance along a smaller number of input or output dimensions. Loudoun County is an example of a generalist county; it performs above average along 7 of the 8 dimensions. In contrast, Amelia is more of a niche county; it only outperforms the average along 4 of the 8 dimensions.[2] Generalist counties represent a more flexible standard, because they are able to remain efficient along a much wider range of input/output weight combinations. Niche counties, on the other hand, are extremely restrictive as reference counties; they remain efficient for a very limited range of weight combinations. In general, one would expect the scores of other counties to drop more significantly when a more limited number of weight combinations is available for their efficiency calculation; we observe an extreme case of this when Amelia County is used as the urban reference county. In fact, the excessive impact of using Amelia County as the reference county indicates a problem with selection of that county as MPS; one could argue that, though it is rated urban, Amelia County has little in common with the other urban counties and that using it as an urban standard is like comparing apples against oranges.

The "niche/generalist" characterization provides a rough guideline about the impact that a particular MPS will have on the remaining efficiency landscape, but a great deal of variation can

[2]Very few counties could occupy a position even more toward the "niche" side of the continuum– i.e. outperform the average along only two dimensions– and still be found to be DEA efficient.

occur within each point along that continuum. The rural counties, Worcester, MD, and Floyd, VA, both outperform the average along 6 of the 8 dimensions, and are therefore moderate generalists in their performance. However, Floyd County clearly has a much more significant impact on the value efficiency ratios of the other counties than does Worcester, MD. Another characterization of counties is useful in explaining this difference.

Counties can be considered "super-achievers" if they experience an extraordinarily high level of performance along one or more input or output dimensions. Such counties also present problems for the other counties when they are used as the reference community. This again leads such counties to place a high weight along those dimensions at which they excel, making them in essence niche counties along those dimensions. It is not surprising for niche counties to also be "super-achievers"– it is in fact that super-achievement along only a few dimensions that enables them to be DEA-efficient. Amelia County is again a case in point; with only .30% of its land area developed, the remaining urban counties, with an average of 12.6% land area developed, simply cannot stack up.

Generalist counties can also be super-achievers, and as that causes them to place a heavy weight on only a few of their above-average dimensions, they in effect turn into niche counties along those few dimensions. Floyd county is an example of such a county. As mentioned earlier, Floyd County is above average on several dimensions, but its concentration of arts and entertainment establishments is a particularly impressive 5.26 establishments per square mile of developed area. The other rural counties, with an average of 1.54 establishments per developed square mile, are at a great disadvantage when this variable is identified as an important component in the output mix. Similarly, its percentage of developed area, at .10%, significantly outperforms the rural average of 1.96%.

Conclusions and Policy Implications

Our results suggest that there are significant differences between rural and urban counties

in the generation of quality of life. The set of most efficient rural counties defines a frontier that is significantly different from that formed by the set of most efficient urban counties. The efficient frontier appears to be a dynamic entity that changes as counties develop, with ability to perform along the different dimensions dependent upon the development history of the county. Not surprisingly, rural counties that are projected onto their efficient frontiers outperform similarly projected urban counties along the environmental dimensions. Urban counties, on the other hand, generally outperform rural counties along all of the socioeconomic dimensions considered, including the index of affordability. This suggests that there has been a certain amount of tradeoff in community development patterns, with counties advancing along the socioeconomic dimensions at the expense of the environmental dimensions.

This analysis is unable to determine whether such tradeoffs are an inevitable result of development, or merely a common, historical result. Our efficiency measurements are based upon observed performances rather than upon an underlying judgment of what production levels, and input/output mixes, are theoretically possible. Among the Mid-Atlantic counties, however, there are no examples of counties that have successfully achieved relatively high levels of performance along several dimensions of their socio-economic quality of life without compromising some performance along the environmental dimensions. This result may change over time as an increased awareness of the environmental impacts of development decisions, as well as the role of the environment in residents' judgment of quality of life, influences the manner in which development occurs and the types of development decisions made.

The economics/environment tradeoff results are not as bleak as they initially appear, however. Further analysis of the results regarding technical efficiency indicate that most of the rural frontier is encompassed by the urban frontier; when a joint frontier is run for the projected data, that frontier is predominantly composed of urban units, and all but nine rural counties become inefficient. Those rural counties that remain efficient are generally counties with exceptional environmental performance. One interpretation of this result is that, in general, rural

counties are not as efficient as urban counties in producing quality of life; i.e. urban counties are generally able to produce more output per unit of input than do rural counties. In just two dimensions, such a situation might appear as shown in figure 2, but with only a small portion of the rural frontier jutting out beyond the urban frontier along the environmental dimension. An alternative way to interpret this figure, however, is that for many levels of environmental quality urban counties are able to achieve greater socio-economic development at the same level of environmental degradation as rural counties.

Together, these results suggest that "pristine" environmental conditions– as measured by extremely high performance along the environmental dimensions– may not be consistent with urbanization and socio-economic development, but that this required tradeoff only holds for very high levels of environmental quality. Below those environmental quality levels, urban counties are able to achieve higher socioeconomic performance than rural counties without further sacrifice in environmental quality. This suggests that the process of urbanization shifts and rotates the efficient frontier in such a way that, below a certain, relatively "pristine" level of environmental quality, urban counties can perform better along the socio-economics dimension while performing at least as well as the rural counties along the environmental dimensions.

DEA can therefore provide a great deal of information about the technical efficiency of rural and urban counties and about the production frontiers that emerge from observed data. When a county is found to be technically inefficient, information from the DEA analysis can be used to identify the shortest route to the efficient frontier. However, this efficiency analysis tells us nothing about the relative desirability of that point on the frontier compared to other points along the efficient frontier. Value efficiency analysis provides added depth to the results by identifying desirable combinations of outputs and using that standard to further distinguish among both efficient and inefficient units.

The results of the value efficiency analysis illustrate rankings for the mid-Atlantic counties based on their "value-efficiency," which measures how much the unit's outputs must be expanded

to reach beyond the efficiency frontier to the isovalue curve associated with the most preferred solution. This analysis considers that even those counties on or near the technical efficiency frontier may not be producing a mix of outputs that is socially desirable, and ranks counties accordingly. In our study, the VEA analysis yielded a broad range of value efficiencies among mid-Atlantic counties in the production of quality of life. This range of efficiencies is highly sensitive to the county that is selected as the reference county. If the reference county has a mix of inputs and outputs and performance levels that are similar to the remaining counties, the value efficiency scores will remain relatively closer to the technical efficiency scores received through DEA. However, if the county chosen as the reference county is in some ways unique or unrepresentative of the other counties, the value efficiency scores of the remaining counties will fall far below their corresponding DEA scores.

DEA and VEA therefore represent flexible methodologies for integrating environmental and socioeconomic variables in an analysis of regional differences in quality of life. These methods provide a great deal of information about relative production performance while requiring few ex ante assumptions about functional relationships or specific input/output weights. The uncommon flexibility of DEA allows plenty of room for innovation in application, with the traditional view of "productivity" expanded to encompass other abstract multidimensional concepts such as quality of life or standard of living. In this study we find the methodologies very useful in providing insight about the distinctions between and among rural and urban counties in generating quality of life, and we believe there is ample opportunity for creative application of such techniques in interdisciplinary research.

References

Allen, R., A. Athanassopoulos, R.G. Dyson, and E. Thanassoulis. 1997. "Weights restrictions and value judgements in data envelopment analysis: Evolution, development and future directions." Annals of Operations Research 73: 13-34.

Banker, Rajiv D. and Richard C. Morey. 1986a. "The use of categorical variables in data envelopment analysis." Management Science 32(12): 1613-1627.

Banker, Rajiv D., and Richard C. Morey. 1986b. "Efficiency analysis for exogenously fixed inputs and outputs." Operations Research 34(4): 513-521.

Blomquist, Glenn C., Mark C. Berger, and John P. Hoehn. 1988. "New estimates of quality of life in urban areas." American Economic Review 78(1): 89-107.

Bowlin, William F. 1998. "Measuring perf ormance: An introduction to data envelopment analysis (DEA)." Journal of Cost Analysis (Fall): 3-27.

Charnes, Abraham, William Cooper, Arie Y. Lewin, and Lawrence M. Seiford. 1994. *Data Envelopment Analysis: Theory, Methodology and Applications.* Kluwer Academic Press, Boston.

Charnes, A., W.W. Cooper and E. Rhodes. 1978. "Measuring efficiency of decision making units." European Journal of Operations Research 2: 429-444.

Cooper, William W., Lawrence M. Seiford, and Kaoru Tone. 1999. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References, and DEA-Solver Software.* Kluwer Academic Press, Boston.

Deller, Steven C., Sue Tsung-Hsiu Tsai, David Marcouiller, and Donald B.K. English. 2001. "The role of amenities and quality of life in rural economic growth." American Journal of Agricultural Economics 83(2): 352-365.

Douglas, S. and H.J. Wall. 1993. "'Voting with your feet' and the quality of life index: A simple non-parametric approach applied to Canada." Economics Letters 42: 229-236.

Gabriel, Stuart A., Joe P. Mattey, and William L. Wascher. 2003. "Compensating differentials and evolution in the quality-of-life among U.S. states." Regional Science and Urban Economics 33: 619-649.

Golany, Boaz and Sten Thore. 1997. "Restricted best practice selection in DEA: An overview with a case study evaluating the socio-economic performance of nations." Annals of Operations Research 73: 117-140.

Halme, Merja, Tarja Joro, Pekka Korhonen, Seppo Salo, and Jyrki Wallenius. 1999. "A value

efficiency approach to incorporating preference information in data envelopment analysis." Management Science 45(1): 103-115.

Hirschberg, J.G., E. Maasoumi, E. and D.J. Slottje. 1998. "The environment and the quality of life in the United States over time," Working Paper 654, Melbourne - Department of Economics.

Korhonen, Pekka, Risto Tainio and Jyrki Wallenius. 2001. "Value efficiency analysis of academic research." European Journal of Operational Research 130: 121-132.

Korhonen, Pekka, Aapo Siljamake, and Margareta Soismaa. 1998. "Practical aspects of value efficiency analysis." International Institute for Applied Systems Analysis, IIASA Interim Report, IR-98-042. 17 pp.

Roback, J. 1982. "Wages, rents, and the quality of life." Journal of Political Economy 90: 1257-1278.

Rosen, S. 1979. Wage-based indexes of urban quality of life, in *Current Issues in Urban Economics*, Mieszkowski, P and M. Straszheim (eds). Johns Hopkins Press, Baltimore.

Tiebout, C. 1956. "A pure theory of local expenditures." Journal of Political Economy 64: 416-424.

Walden, John B. and James E. Kirkley. 2000. "Measuring technical efficiency and capacity in fisheries by data envelopment analysis using the General Algebraic Modeling System (GAMS): A workbook." NOAA Technical Memorandum NMFS-NE-160. 15 pp.

Table 1: Variables included in Quality of Life model	
Environmental dimension	EPA's cancer risk index (input) % of land area developed (input)
Social dimension	Teacher/pupil ratio (input) % of population 25 and older who are high school graduates (output) # of arts, recreation, and entertainment establishments per developed square mile (output)
Economic dimension	Median household income (output) % of population below poverty level (input)
Non-discretionary amenity variables	Amenity index (output)

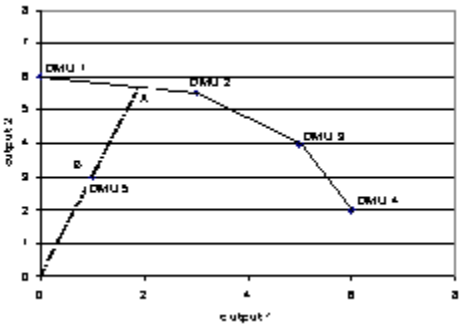

Figure 1: Efficient frontier with 5 production units.

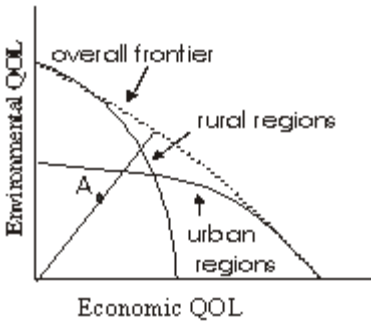

Figure 2: Determining the appropriate frontier

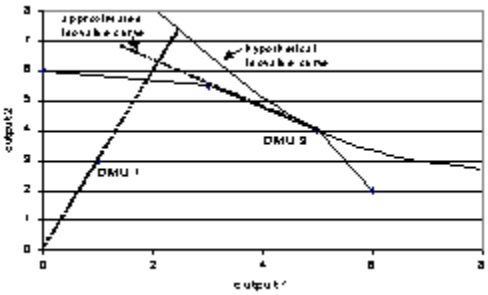

Figure 3: Value efficiency analysis of decision-making units.

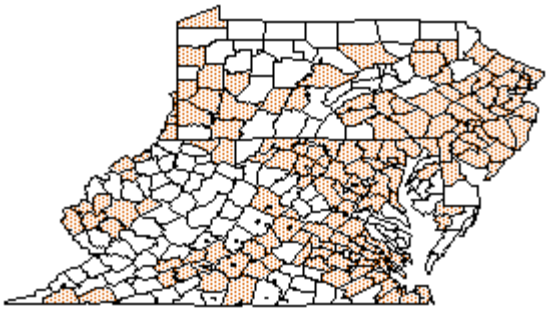

Figure 4: Rural/urban distinctions are determined by Beale codes. Urban counties are shaded.

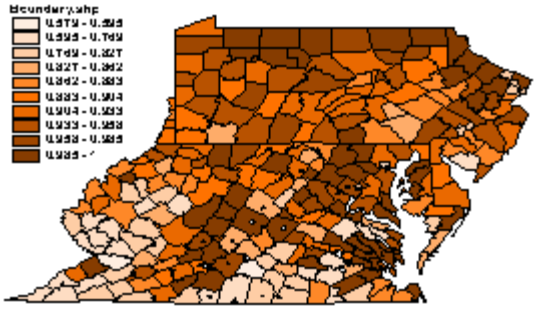

Figure 5: Technical efficiency results derived from data envelopment analysis.

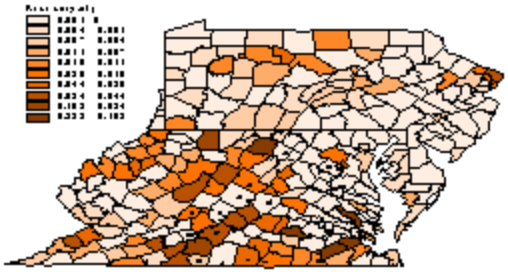

Figure 6: Change in efficiency score when value efficiency analysis is used to compare counties to reference communities rather than data envelopment analysis.

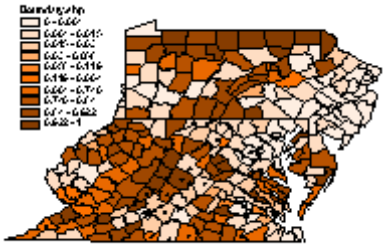

Figure 9: Value efficiency scores when Amelia County, VA, is used as the reference county for urban counties and Floyd County, VA, is used as the reference county among the rural counties.