Valuing Community Attributes in Rural Counties of West Virginia Using Data Envelopment Analysis

Maribel N. Mojica
Alcorn State University
Alcorn State, MS, USA
mmojica@alcorn.edu

Tesfa G. Gebremedhin
West Virginia University
Morgantown, WV, USA
tesfa.gebremedhin@mail.wvu.edu

Peter V. Schaeffer
West Virginia University
Morgantown, WV, USA
peter.schaeffer@mail.wvu.edu

Abstract

As quality of life measures are related to increased economic activity, it becomes increasingly important to develop indicators as accurate measures of the well-being of the residents in a community. This study use Data Envelopment Analysis (DEA) to analyze community attributes of rural counties in West Virginia using variables determining quality of life. County level data is used to identify counties that are inefficient as measured in terms of socioeconomic factors. Desirable community attributes such as employment, median household income, median house value, health index, number of personal care establishments, and number of high school graduates were used as output variables. Input variables representing the undesirable characteristics of counties include population density, unemployment rate, per capita tax, number of persons below poverty, and crime rate. The analysis seeks to determine efficiency levels in the rural areas of the State. The results show that majority of the rural counties in the State lie on the efficiency frontier, while others are classified to be inefficient. The research findings that can be used as indicators of community performance and to evaluate counties in terms of creating quality of life are of interest to policy makers.

Keywords: data envelopment analysis, community attributes, output variables, input variables, efficiency levels, quality of life, principal components analysis

---

1.0 Introduction

West Virginia is a state in transition, with some regions experiencing rapid economic growth and others remaining predominantly rural and impoverished. Rural mining regions, particularly in the central and southern part of the State, are experiencing slow economic growth (Bukenya, Gebremedhin, & Schaeffer, 2003). In stark contrast to most of the rest of the U.S. economy, West Virginia experiences high unemployment, poverty declining economy and out migration (Deavers & Hope, 1992). The State also ranks among the last in the nation in income and wealth

(Dilger & Witt, 1994; Haynes, 1997). West Virginia's unique position as a state in transition offers the opportunity to evaluate the quality of life in the state.

Quality of life as a concept can mean different things to different people, encompassing such notions as "well-being" centered on the individual to a "good place" centered on the location. Analyzing the quality of life in residential areas has gained increasing attention in recent studies as it becomes important to understand the value of the social and economic characteristics associated in different communities as measures of quality of life. Several studies (Blomquist, Berger, & Hoehn, 1988; Deller, Tsai, Marcouiller, & English, 2001; Gabriel, Mattey, & Wascher, 2003; Roback, 1982; Rosen, 1979) have attempted to measure and quantitatively analyze the concept of quality of life to determine its importance in economic growth and development. As quality of life measures are related to increased economic activity, it becomes increasingly important to develop indicators as accurate measures of the well-being of the residents in a community. As the concept of quality of life is multi-dimensional, it becomes challenging to develop quantitative measures determining quality of life valuations.

Furthermore, quality of life affects the decision patterns of individuals in choosing residential areas which increases the importance of understanding the valuation of attributes in different communities. This is based on the hypothesis that residents will "vote with their feet" (Tiebout, 1956) and will therefore move to communities with higher quality of life. The value of social and economic characteristics of communities determines quality of life and affects migration patterns as it becomes a significant factor to attract people in certain locations.

In this study, Data Envelopment Analysis (DEA) is used to analyze the efficiency of communities in producing quality of life using the rural counties of West Virginia as decision making units. The objective of the study is to evaluate the performance of rural communities in West Virginia in providing quality of life using socio-economic factors. Data Envelopment Analysis was used to calculate efficiency scores among rural counties in the State as quantitative measures of the efficient production of quality of life within communities. The analysis is focused on rural counties as it becomes increasingly important to consider the open spaces, the natural amenities and the values given to smaller towns in providing quality of life.

2.0 Literature Review

Several studies have used different sets of variables reflecting social, economic, and environmental factors to evaluate the quality of life in certain communities. Economists have attempted to estimate the outcome of observed individual behavior in understanding how the society's well being is affected by location attributes. Rosen (1979) studied how quality of life variables affect location decisions, estimated its market price and provided city rankings in terms of the city's attractiveness as reflected by measures of quality of life. The study suggests that differentials in wage and rent can be explained by the characteristics of the city. The analysis was done by using quality of life indicators such as level of pollution, type of climate, unemployment rate and population growth.

Blomquist et al. (1988) used housing expenditure and wage differentials to explain differences in quality of life. Quality of life rankings were provided by constructing indices using preference-based weights from hedonic estimation. To rank communities based on quality of life indices, various amenity variables such as climatic, environmental, and urban conditions were used to estimate the hedonic equations. The results show that there is significant compensation for location-specific and non-traded amenities in the labor and housing market. The study presented a comparison within and among 253 cities in the US and the conclusions support the argument that quality of life is an important factor considered in location decisions. An extension of the study by Gabriel et al. (2003) analyzed the changes in quality of life rankings over time for United States. In addition to housing expenditure and wage differentials, the study included nonland cost of living in the hedonic estimation to reflect the capitalization of amenities. The results revealed that some states recorded a substantial deterioration in quality of life predominantly due to limited infrastructure investment. Other states have improved quality of life while some other states remained the same. Giannias (1998) also used housing and city characteristics as indicators to rank Canadian cities according to their attractiveness in terms of quality of life.

Deller et al. (2001) used five indices designed to capture specific amenity and quality of life characteristics. One of the objectives of the study is to construct measures of amenities and quality of life among non-metropolitan counties in the United States. This was done by using principal components analysis where groups of variables are condensed into a single scalar measure that captures the information in the original data. Five amenity attributes were constructed to represent indicators of quality of life in the rural areas. Of the five amenity attributes, the empirical findings showed that amenity attributes were positively related to at least one measure of economic growth reflected by the change in population, employment, and per capita income.

Kahn (1995) used data from the 1980 and 1990 Census of Population and Housing to rank the cities of Chicago, Houston, Los Angeles, San Francisco, and New York based on quality of life. The method of ranking used the differences in skills across cities as well as wage estimates. It is based on a revealed preference approach where a person’s characteristics are evaluated at each city’s estimated price vector. This approach allowed the ranking of the cities even if no attributes of the cities are observed, with the assumption that all agents in the economy have similar preferences. If there is zero cost of migration and all agents have equal skills, the equilibrium consists of differences in rental and wages, and people will be distributed across cities such that no person can move to another city to raise his utility. The analysis was done using wage and rental regressions which resulted in rankings that Los Angeles and San Francisco have higher quality of life than Chicago and Houston in 1980 and 1990. In addition, Kahn found that quality of life in New York fell during the 1980s.

Furthermore, in urban economic literature, Douglas (1997) found that living standards were highest in the Northwest in the years 1970 and the 1980, while in 1990 it is highest in the south Atlantic coast. The study is based on the measurements of relative attractiveness of areas by observing individual location decisions. A random utility model was used to derive a measure of relative standard of living across U.S. states through pairwise comparisons which are then used to rank states. Standard of living was used in the study to refer to the opportunity to obtain a higher real income as an expanded definition of quality of life in the literature. Data on migration rates are analyzed to construct standard of living indices and rankings of the U. S. states which are found to be significantly affected by population. The results also indicated the effect of energy market dislocations and other economic shifts in the 1970s on population distribution. As an extension, Wall (2001) reexamined the use of migration rates to estimate compensating differentials as indicators of regional quality of life in the United Kingdom. The study adds to the existing literature by relaxing the assumption of the independence of moving costs with the direction of the move, separating labor market conditions from other amenities, and controlling for the effect of contiguity. The results show the high correlation of the regional rankings of quality of life and standard of living in the UK.

3.0 Theoretical Background

Data envelopment analysis (DEA) is a mathematical programming technique used to measure efficiencies of decision making units (DMUs). The technique was introduced by Charnes, Cooper, and Rhodes (1978) and is widely used as a non-parametric approach originally developed to analyze the performance of different organizations. DEA is an approach extensively used by different sectors in management science, agriculture, and economics. For example, Speelman, D'Haese, Buysse, & D'Haese (2007) used Data Envelopment Analysis to measure the efficiency of water use in small scale irrigation schemes; Fogarasi and Latruffe (2007) used DEA to estimate the efficiency of crop and dairy farms; Gorton and Davidova (2004) used the technique in analyzing farm productivity; and Oude and Silva (2003) employed DEA to analyze the efficiencies of different heating technologies.

In this study, DEA is used to analyze efficiency levels of counties by maximizing the desirable attributes such as employment, median household income and health index, while minimizing undesirable attributes such as unemployment rate and crime rate. The reasonable assumption is that communities want to maximize the desirable attributes and minimize the undesirable community attributes. The method provides a non-parametric approach in ranking communities and to analyze the contributions of different socio-economic factors in producing quality of life.

Evaluation of relative efficiency of each unit (county) is done by using the ratio between weighted outputs (desirable factors) over the weighted inputs (undesirable factors), which is the most basic DEA formulation (Anderson, Sweeney, & Williams, 2005). A unit (a county, in this case) with an efficiency score of 1 is considered efficient while a ratio of less than one indicates inefficiency relative to other units. This allows DEA to provide a way to measure efficiencies while not requiring the specification of the production relationships between inputs and outputs (Marshall & Shortle, 2005). It is applicable when using different variables of differing units and is able to accommodate multiple units and outputs.

4.0 Empirical Model

Data envelopment analysis (DEA) is based on the concept that if a decision making unit (a county in this case) is using less input than another unit to produce the same level of output, that unit can be considered as more efficient based on an observed efficiency measure. Efficient counties lie on the efficiency frontier while the inefficient ones lie outside the frontier. The frontier is represented by the composite county.

Since DEA is a non-parametric method, it is not necessary to have a direct relationship between the inputs and outputs considered in the analysis (Marshall et al., 2005). Counties that are relatively efficient will produce relatively more desirable attributes (outputs) per unit of undesirable attributes (inputs) than counties that are relatively inefficient.

The model is based on a study by Speelman et al. (2007) and the sample problem presented by Anderson et al. The model that follows below is for a data set on M inputs (undesirable attributes) and N outputs (desirable attributes) for each of the J rural counties in West Virginia:

$$\textbf{Min } \theta = \alpha_{i1} X_{i1} + \ldots \alpha_{im} X_{im} \qquad (1)$$

Subject to:

$$-y_i + Y\pi \geq 0, \qquad (2)$$

$$-\theta x_i + X\pi \leq 0, \qquad (3)$$

$$J1'\pi = 1 \qquad (4)$$

$$\theta, \pi \geq 0 \qquad (5)$$

Where θ is an efficiency measure with a value between zero and one, one indicating that the county is relatively efficient. For the i-th county, input and output data are presented by column vectors x_i and y_i , respectively, while X and Y are the input and output matrices representing the data for all J counties. J1 is a Jx1 vector of ones and π is a Jx1 vector of constants. The objective function is to minimize θ which is the same as minimizing the undesirable community attributes (inputs). The DEA programming model is composed of thirty-four decision variables and twelve constraints, excluding the non-negativity constraint.

5.0 Data Types and Sources

County-level data from USDA Economic Research Service (ERS) and the Geospatial and Statistical Data Center was used to analyze factor inputs and outputs (USDA-ERS, 2007; Geostat, 2007). The inclusion of variables is determined by the availability of information for all units (rural counties) in consideration as well as for variable correlations. Variable correlations will not affect the analysis but are used to avoid inclusion of many variables measuring the same community attribute. Data search resulted in the inclusion of output variables: employment, median household income, median house value, health index, personal care facilities, and high school graduates. These output variables are the development factors that the counties would maximize. Input variables include population density, unemployment rate, per capita tax, number of persons in poverty, and crime rate. These are the input variables the counties would minimize. The output and input variables included in the model are presented in Table 1.

Following Marshall et al. (2005), rural counties were defined using the ERS rural-urban continuum codes called Beale Codes. Counties with Beale Code of 0-3 were considered urban, while those with 4-9 were considered rural (Figure 1).

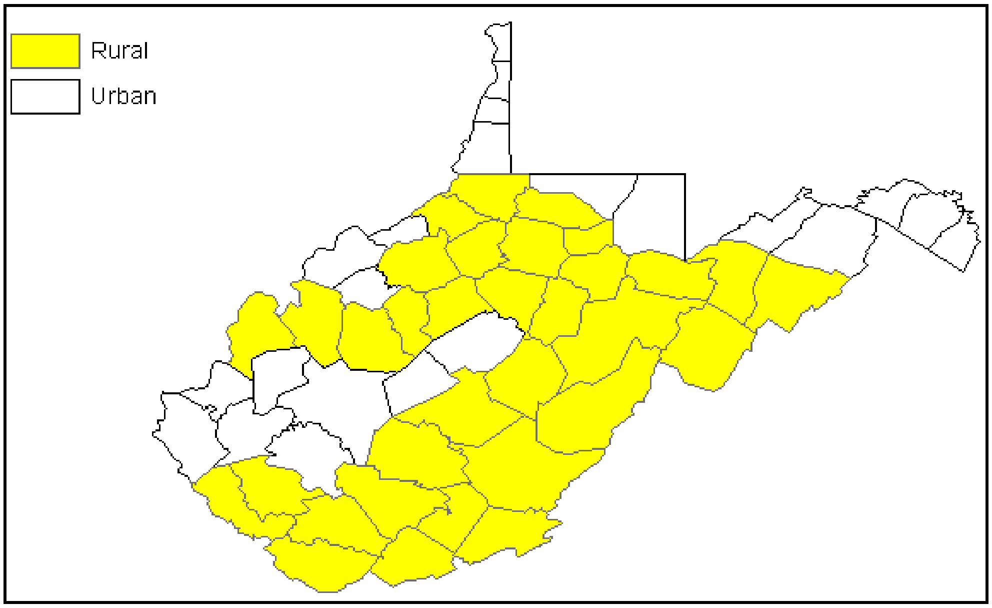

Figure 1. Rural/Urban Distinctions in West Virginia Counties

Table 1. *Descriptive Statistics of Outputs and Inputs Used in DEA*

	Unit	Average	Standard Deviation	Minimum	Maximum
OUTPUTS					
Employment (EMP)	number	5503.45	5573.53	336.00	21998.00
Median Household Income (MHY)	dollars	18000.06	2268.75	13141.00	21655.00
Median House Value (MHV)	dollars	40593.94	7463.16	15800.00	51600.00
Health Index (HLTH)	number/ points	8.68	2.61	4.20	15.50
Personal Care Facilities/ Establishments (PERFAC)	number	3.64	3.87	0.00	17.00
High School Graduate 25 years and older (HIGRAD)	number	10063.30	8399.16	2671.00	32783.00
INPUTS					
Population Density (POPDEN)	persons per square mile	55.88	44.90	9.58	184.67
Unemployment Rate (UNEMP)	percent	13.18	3.72	7.70	22.40
Per Capita Tax (PCTAX)	number	281.71	84.61	90.68	416.90
Number of people below poverty (BELOPOV)	number	4836.82	3599.94	1081.00	13852.00
Crime rate (CRIME)	number	1320.85	614.34	479.00	2944.00

6.0 Empirical Results and Analysis

The model presented above is estimated using Microsoft Excel Solver (Frontline Systems, 2007). Table 2 shows the distribution of efficiency scores of the counties with 6% having an efficiency score between 0.61 and 0.70, 12% with scores between 0.81 and 0.90, 21% between 0.91 and 0.99, and 61% have an efficiency index of 1. Counties with efficiency scores of less than one are considered inefficient, while those with a score of one are considered efficient. The average overall efficiency for all counties is 0.92.

Table 2. *Efficiency Estimates of Rural Counties in West Virginia*

Efficiency Score	Number of Counties	% of Counties
0.00 - 0.60	0	0
0.61 - 0.70	2	6
0.71 - 0.80	4	12
0.81 - 0.90	7	21
0.91 - 0.99	0	0
1.00	20	61
Average Score: 0.92		

Figure 2 further shows the efficiency scores of rural counties in West Virginia. Most of the counties (61%) have an efficiency score of one and are therefore considered to be relatively efficient. This means that these counties are maximizing the desirable community attributes and minimizing the undesirable attributes as much as the composite county. These are the counties that lie on the efficiency frontier. Twenty-one percent of the units have an efficiency score of between 0.81 and 0.90 which is a relatively high score of efficiency. The values mean that the composite county is more efficient than these counties and are relatively inefficient. The score indicates that the composite county can have at least the same level of each output that the counties with efficiency scores of between 0.81 and 0.90 can obtain but with fewer amounts of inputs. That is, the composite county can have the same level of desirable characteristics (outputs) as the inefficient counties but with less undesirable attributes (inputs), which is no more than 81 to 90% of the inputs by the inefficient counties. Two counties, Roane and Wyoming, have the lowest efficiency scores of 0.67 and 0.70 indicating that the composite county can have the same level of desirable characteristics but with only 67 and 70% of Roane and Wyoming's undesirable characteristics.

The results of the DEA analysis also showed additional information on the slack/surplus variables for each county (Table 3). The figures further support the efficiency scores of the counties and their interpretations. For example, in Table 3, the slack/surplus estimation for Roane County shows that the composite county has at least as much of each level of output as Roane County has but with 385 more employment, $2900 more of median household income, $7754 more house value, and 0.95 unit more of personal care establishment. The slack with zero values for the input constraints indicate that the composite county has only 67% of the level of population density, per capita tax, and number of persons below poverty of Roane County. These results further support that the counties with an efficiency score of less than 1 are inefficient, i.e. Roane County.

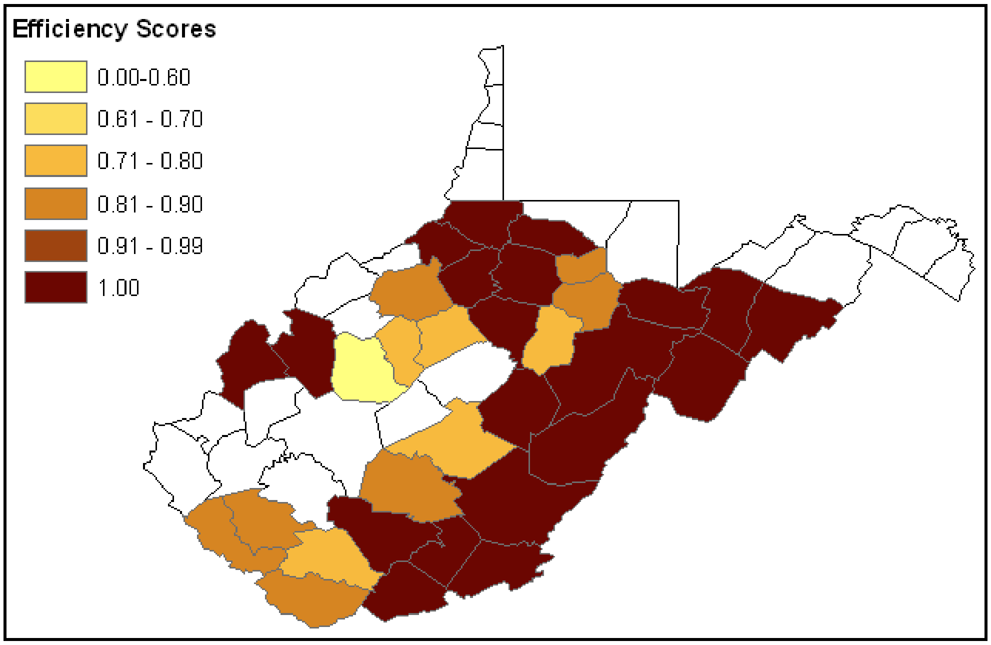

Figure 2. Efficiency Scores of Rural Counties in West Virgina

Table 3. *Slack values for Roane County*

	Unit	**Slack**
Employment (EMP)	Number	385.32
Median Household Income (MHY)	Dollars	2900.88
Median House Value (MHV)	Dollars	7754.79
Health Index (HLTH)	number/points	0.00
Personal Care Facilities/Establishments (PERFAC)	Number	0.9468
High School Graduate 25 years and older (HIGRAD)	Number	0.00
Population Density (POPDEN)	persons per square mile	0.00
Unemployment Rate (UNEMP)	Percent	0.7715
Per Capita Tax (PCTAX)	Number	0.00
Number of people below poverty (BELOPOV)	Number	0.00
Crime rate (CRIME)	Number	46.07

7.0 Conclusions

The results suggest that the majority of the rural counties in West Virginia are relatively efficient in increasing quality of life. Since efficiency measures are based on community performances and not on production levels, it is difficult to determine the reasons for such efficiency values. However, the rural counties in the Eastern part of West Virginia were observed to be more efficient in generating quality of life relative to the other rural counties in the State. This may be due to the fact that counties that were observed to be efficient are in close proximity to the more developed counties.

The results of DEA presented a reliable method of measuring the efficiency of rural counties as indicators of quality of life based on available socio-economic data. These are useful information as most people base their migration decisions on quality of life in a community. The analysis also provides information of interest to policy makers as an indication of community performance which can be used for evaluating quality of life of counties and initiating development programs. In addition, the counties that are found to be efficient can act as models for efficiency to the inefficient counties, illustrating the achievable levels of outputs given the characteristics of the counties.

Data envelopment analysis is a flexible method of analyzing community attributes using socio-economic variables. It is a beneficial method in measuring efficiency without the necessity of having direct relationships between output and input variables. DEA is a very useful method in determining the differences between rural counties in West Virginia about generating quality of life.

8.0 References

Anderson, A. R., Sweeney, D. J., & Williams, T. A. (2005). *An Introduction to Management Science: Quantitative Approaches to Decision Making* (10th ed.). Mason, Ohio: South-Western, a division of Thomson.

Blomquist, G. C., Berger, M. C., & Hoehn, J. P. (1988). New Estimates of Quality of Life in Urban Areas. *American Economic Review, 78*, 89-107.

Bukenya, J. O., Gebremedhin, T. G., & Schaeffer. P. V. (2003). Analysis of Quality of Life and Rural Development: Evidence from West Virginia Data. *Growth and Change, 34*(2), 202-218.

Charnes, A., Cooper, W., & Rhodes, E. (1978). Measuring the Efficiency of Decision Making Units. *European Journal of Operational Research*. *2*, 429-444.

Deavers, K. L. and R. A. Hope. (1992). Overview of the Rural Poor in the 1980s. In C. Duncan (Ed.). *Rural Poverty in America* (pp. 3-20). Westport, CT: Greenwood Publishing Group, Inc.

Deller, S. C., Tsai, T., Marcouiller, D. W., & English, D. (2001). The Role of Amenities and Quality of Life in Rural Economic Growth. *American Agricultural Economics Association*, *83*(2), 352-365.

Dilger, R. J. and T. M. Witt. (1994). West Virginia's Economic Future. In D. R. Jay & T. S. Witt (Eds.). *West Virginia in the 1990s: Opportunities for Economic Progress* (pp. 3-15). Morgantown, WV: West Virginia University Press.

Douglas, S. (1997). Estimating Relative Standard of Living in the United States Using Cross-Migration Data. *Journal of Regional Science*. *37*(3), 411-436.

Fogarasi, J. & Latruffe, L. (2007). Technical Efficiency and Technology in Eastern and Western Agriculture: A Comparison of Crop and Dairy Farms in Hungary and France. Paper presented at the joint IAAE – 104th EAAE Seminar. Budapest, Hungary, Sept. 6-8, 2007.

Frontline Systems Inc. (2007). Retrieved October 21, 2007, from http://www.solver.com/contact.html

Gabriel, S. A., Mattey, J. P., & Wascher, W. L. (2003). Compensating Differentials and Evolution in the Quality of Life among US States. *Regional and Urban Science Economics*, *33*(2003), 619-649.

Geostat. Geospatial and Statistical Data Center. University of Virginia Library. (2007). Retrieved October 21, 2007, from http://fisher.lib.virginia.edu/cgi-local/ccdbbin/countysort2.cgi

Giannias, D. (1998). A Quality of Life Based Ranking of Canadian Cities. *Urban Studies*, *35*(12), 2241-2251.

Gorton M. & Davidova, S. (2004). Farm Productivity and Efficiency in the CEE Applicant Countries: A Synthesis of Results. *Agricultural Economics, 30*, 1-16.

Hayness, A. F. (1997). *Poverty in Central Appalachia: Underdevelopment and Exploitation*. New York, NY: Garland Publishing, Inc.

Kahn. M. E. (1995). A Revealed Preference Approach to Ranking City Quality of Life. *Journal of Urban Economics*, *38*, 221-235.

Marshall, E. & Shortle, J. (2005). Using DEA and VEA to Evaluate Quality of Life in the Mid-Atlantic States. *Agricultural and Resource Economics Review*, *34*(2), 185-203.

Oude L. & Silva, E. (2003). CO_2 and Energy Efficiency of Different Heating Technologies in Dutch Glasshouse Industry. *Environmental and Resource Economics*, *24*, 395-407.

Roback, J. (1982). Wages, Rents, and the Quality of Life. *Journal of Political Economy*, *90*(6), 1257-1278.

Rosen, S. (1979). Wage-based Indexes of Urban Quality of Life. *Current Issues in Urban Economics*. Baltimore: Johns Hopkins University Press.

Speelman, S., D'Haese, M., Buysse, J., & D'Haese, L. (2007). Technical Efficiency of Water Use and its Determinants, Study at Small-Scale Irrigation Schemes in North-West Province, South Africa. Paper presented at the 106th seminar of EAAE. Montpellier, France, Oct. 25-27, 2007.

Tiebout, C. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, *64*(5), 416-424.

USDA-ERS. U.S Department of Agriculture-Economic Research Service. (2007). Profiles of America. Retrieved October 21, 2007, from http://maps.ers.usda.gov/profiles/webcensusdownload.aspx

Wall, H. J. (2001). Voting with your feet in the United Kingdom: Using cross-migration rates to estimate relative living standards. *Papers in Regional Science*, *80*(1), 1-23.