ANALYSIS OF RURAL QUALITY OF LIFE AND HEALTH: A SPATIAL APPROACH

James O. Bukenya, Tesfa G. Gebremedhin and Peter V. Schaeffer[1]

Abstract

This paper examines the relationship between quality of life, health and several socioeconomic variables. The analysis utilizes empirical data obtained from a survey questionnaire administered on a random sample of over 2000 residents in twenty-one counties in West Virginia, and spatial data obtained by geocoding the survey respondents' addressees. Quality of life is measured by a three-point categorical measure of overall satisfaction and an ordered probit model is used to examine the relationships. The empirical results are consistent with the theoretical predictions and indicate, for instance, that quality of life satisfaction increases with income and education while it decreases with unemployment.

Key Words: Quality of life, health, rural, development, spatial analysis

Classification JEL: C8, D60, 131, 112

[1] Former Graduate Research Assistant, Professor of Agricultural and Resource Economics and Director/Professor, Division of Resource Management, West Virginia University. The study was supported by the Hatch funds appropriated to West Virginia University Agriculture and Forestry Experiment Station. Selected paper to be presented at the American Agricultural Economics Association annual meeting, Long Beach, California, July 28 – 31, 2002

ANALYSIS OF RURAL QUALITY OF LIFE AND HEALTH: A SPATIAL APPROACH

Introduction

In the recent years the notion of quality of life has attracted much attention (Clark and Oswald, 1994; Heubusch, 1998; Reichert and Rudzitis, 1992; Easterly, 1999; Dissart and Deller, 2000; Deller et al., 2001; Hall and Jones, 1997, 1998). As a concept, quality of life can mean different things to different people, encompassing such notions as "well-being" centered on the individual to "good place" centered on the location. Regardless of which meaning is considered, the notion of quality of life has considerable implications for planning (Dissart and Deller, 2000). Recent studies have evaluated the quality of life within local jurisdictions, and among nations (Mencken, 1998; Gerdtham et al., 1997; Sousa-Poza et al., 2000). The methods of analysis range from journalistic approaches to elaborate scientifically based research strategies. Despite the burgeoning literature, there is little unanimity on this subject.

To encapsulate its full meaning, it is not sufficient to consider only the process of provision of, and access to, a better environment and better facilities. Rather, quality of life should also include a consumer-oriented perspective that is concerned with the manner of delivery of goods, services or facilities and with the quality of the environment and, finally, with the experience that arises from the consumption of the goods or services. From a policy perspective, quality of life measurements play an important role of allowing the identification and evaluation of those factors that have the greatest impact on society's well being and are amenable to improvement.

Utilizing results previously published, we reinvestigate the effects of factors that have been found to influence a "satisfied" or "dissatisfied" response to questions about

quality of life, well being or happiness. This is done in an attempt to reconcile inconsistencies and conflicts in previous quality of life studies. Unlike previous studies, the current paper addresses the geographical differences across the studied regions through spatial analysis. Over the last 15 years conceptual models of place well being in sociology have continued to emphasize the importance of regions and geographical space (Lyson and Tolbert, 1996). This is particularly true in research with politically constructed geographical units of analysis, such as counties (Land and Deane, 1992; Doreian, 1981; Mencken, 1998).

Previous research (Killian and Tolbert, 1993) shows that there is considerable work-related commuting across county borders, thus supporting the argument that the expansion and contraction of economic activity is impervious to politically constructed geographical borders (Anselin, 1988; Land and Deane, 1992). For example, Mencken (1997) reported higher satisfaction levels of well being in the southern Appalachian counties at the end of the 1980s. When he reanalyzed the data and corrected for spatial dependency[2], some of the results changed (Mencken, 1998). In this paper, we try to correct for spatial lag using spatial information obtained from the geocoded survey respondents' addresses[3].

[2] Spatial dependence exists if either the dependent variables or the error terms are correlated with each other. If the dependent variables in the analysis are spatially dependent, then spatial lag is present. If the model does not correct for the lag, regression estimates will be biased. If spatial error is present, the regression estimator will be inefficient. The presence of either spatial lag or spatial error (or both) could therefore substantially change the conclusions of the analysis.

[3] Testing and correcting for spatial dependence in discrete choice models is technically more demanding than treated in this paper (see Fleming, 2002; Smith and LeSage, 2001; Anselin, 2001). However, the approach used here (geocoding respondents' addresses to develop spatial weight matrixes) is deemed sufficient to capture the geographical aspects across the study regions while at the same time examining the effect that residing close to a medical facility may have on one's satisfaction with the quality of life they lead.

The paper has two major objectives. First, to examine the quality of life aspects that are important to and that result from living and working in rural areas. The second objective is to reinvestigate the effects of those factors that have been found in previous studies to facilitate satisfied or dissatisfied responses, and to compare the differences and similarities with previous quality of life studies. The remainder of the paper is organized as follows. The proceeding section outlines the theoretical model upon which the analysis is based. The model is adopted from Gerdtham and Johannesson's (1997) study examining the relationships between happiness, health and economic factors. The section is followed by a discussion of the estimation model, data sources and definitions. The last section present the empirical results followed by the summary and conclusions.

Theoretical Model

Apparently, economic theory emphasizes the process by which individuals rationally allocate resources to meet their needs, thereby "producing" utility. Economists characterize the problem of resource allocation in terms of maximizing utility; subject to constraints on the availability and substitutability of resources (Schuessler and Fisher, 1985). The theory posits that individuals derive utility according to the following function:

$$U_i = U_i(h_i, x_i, s_i) \qquad \text{where,} \tag{1}$$

U_i represents the utility of individual i (i=1, …, I), h_i is the health status of individual i, x_i is a vector of private goods consumed, and s_i is a vector of socioeconomic and spatial variables that affect utility (Gardtham and Johannesson, 1997).

Health is produced according to the following health production function:

$$h_i = f(m_i, h_{0i}, s_i) \qquad \text{where,} \tag{2}$$

m_i is a vector of health goods such as medical care, and h_{0i} is the initial (given) health status. Utility is maximized subject to the budget constraint

$$Y_i = P_{xi}x_i + P_{mi}m_i \qquad \text{where,} \qquad (3)$$

Y_i represents the exogenously given income of individual *i*, P_{xi} is a vector of private goods prices faced by individual *i*, and P_{mi} is a vector of health goods prices faced by individual *i*. This leads to the following indirect utility function:

$$V_i = V_i\left(h_i\left(Y_i, P_{xi}, P_{mi}, h_{0i}, S_i\right), Y_i, P_{xi}, S_i\right) \qquad \text{where,} \qquad (4)$$

V_i represents utility or life satisfaction. The exogenous variables $\left(Y_i, P_{xi}, S_i\right)$ may influence utility either directly or indirectly through the intervening health variable. This suggests two alternative approaches for estimating the utility model. The first approach is to model the intervening health variable explicitly in the following equation system:

$$V_i = \beta_1 + \beta_2 Y_i + \beta_3 P_{xi} + \beta_4 h_i + \beta_5 S_i + \varepsilon_1 \qquad (5)$$

$$h_i = \beta_6 + \beta_7 Y_i + \beta_8 P_{xi} + \beta_9 P_{mi} + \beta_{10} h_{0i} + \beta_{11} S_i + \varepsilon_2 \qquad (6)$$

Alternatively, equation (6) can be substituted for h_i in equation (5) and the following reduced form model estimated:

$$V_i = \beta_{12} + \beta_{13} Y_i + \beta_{14} P_{xi} + \beta_{15} P_{mi} + \beta_{16} h_{0i} + \beta_{17} S_i + \varepsilon_3 \qquad (7)$$

In the above models, β_1 through β_{17} represents coefficients to be estimated and ε_1 through ε_3 are error terms, assumed to have a zero mean and constant variance. It is also assumed that $\text{cov}(\varepsilon_1,\varepsilon_2)=0$. The full structural approach of equations (5) and (6) distinguishes between the indirect effects of the exogenous variables working through health and the direct effects of the exogenous variables, after controlling for health. That is, the model identifies the process underlying the effects of the exogenous variables. The

second approach (equation 7) captures only the total (direct and indirect) effects of the exogenous variables in a reduced form equation. In this paper both approaches are used to evaluate both the direct and indirect effects of the variables.

Estimation Model

Ordered probit or logit models are the appropriate techniques for relationships involving ordinal dependent variables (Hanushek and Jackson, 1977; Amemiya, 1981; Cameron and Trivedi, 1986; Greene, 1993). These techniques take the ceiling and floor effects into account and avoid the use of subjectively chosen scores assigned to the categories. An ordered probit model is used because the dependent variables are ordered responses.

Let V_i^* be a continuous, latent variable representing, for instance, the cardinal utility function of the individual. Linear dependence is assumed between the latent variable V_i^* and X_i, and β and ε_i:

$$V_i^* = \beta X_i + \varepsilon_i, \quad \varepsilon_i \sim \mathrm{N}(0, \sigma^2), \tag{8}$$

The variable V_i^* defines a variable v_i, related to the above-mentioned categories in the following way:

$$v_i = \begin{cases} 0 \text{ if } V_i^* \le \theta_0 \\ 1 \text{ if } \theta_0 < V_i^* \le \theta_1 \\ 2 \text{ if } \theta_1 < V_i^* \end{cases} \tag{9}$$

θ_i = 0,1, are unobservable thresholds. Denoting the cumulative density function of the standard normal distribution as $\Phi(\cdot)$, it follows that the probabilities of an individual belonging to a given category are:

$$\text{Prob}[Y_i = 0] = \Phi[\mu_0 - \alpha X],$$

$$\text{Prob}[Y_i = 1] = \Phi[\mu_1 - \alpha X] - \Phi[\mu_0 - \alpha X], \tag{10}$$

$$\text{Prob}[Y_i = 2] = 1 - \Phi[\mu_1 - \alpha X],$$

with $\alpha = \beta/\sigma$ and $\theta_j / \sigma = 0,1$. Note that only the ratios β/σ and θ_j / σ can be estimated (Dustman, 1996). If the regression model contains a constant term, the full set of coefficients is not identified. A common normalization is to set $\mu_0 = 0$, which means that the estimated coefficients μ_i, i=1 represent the differences in the respective thresholds: $\mu_i = \mu_i - \mu_{i-1}$ (Greene, 1995; Dustman, 1996).

To correct for misspecification, a multiplicative heteroscedasticity ordered probit model is used to estimate the structural equations (5) and (6) and the reduced form equation (7). Equations (5) and (6) are recursive (triangular) systems with a diagonal $\sum$ matrix. That is, there is a unidirectional dependency among health and utility, and the disturbances across equations are assumed contemporaneously. The equations in the utility-health system are thus estimated separately using the ordered probit model (Greene, 1993; Gardtham and Johannesson, 1997).

Data

The micro level data are obtained from a quality of life mail survey conducted in year 2000[4]. The survey was sent to 2000 residents in 21 counties located in the southern and eastern panhandle regions in West Virginia. The study regions are chosen as representative samples of the poor and rich regions in the state. The two regions differ

[4] Completed questionnaires were received from 1,060 individuals (return rate 53 percent), of whom 532 were female and 528 were male. Of the completed survey questionnaires, thirty-two (32) were discarded. The data used in the empirical analysis are based on 1,028 questionnaires (return rate 51.4 percent).

somewhat by their proximity to big metropolitan areas (Baltimore and Washington DC) in the eastern panhandle region. The residents were selected randomly using telephone numbers. Secondary data are obtained from the Bureau of Business and Economic Research (BBER, 2000), and the Regional Economic Information Systems (REIS, 1998).

To generate the spatial data, the regions of analysis are defined as counties in West Virginia. Topological information is generated through geographical information systems (GIS). GIS is also used to provide the basic measure of spatial linkages (Fortheringham and Wegener, 2000). To create the spatial weight matrix that describes the linkages, the "address geocoding process" in ArcView 3.2 is used. This process creates a theme based on addresses in a table, using a reference feature theme. The "reference theme" (street theme with address ranges on each street segment) is drawn from a TIGER (Topologically Integrated Geographic Encoding and Referencing) file, and the information is converted to an ARC INFO street coverage with address ranges in its attribute table. The data are obtained from ESRI-ArcData online (ESRI, 2000).

To define the spatial variable, the issue of whether proximity to a hospital adds to satisfaction is addressed. Economic theory suggests that by virtue of better access, households located closer to a hospital should have a higher health (utility) bundle. Using this argument, a spatial weight matrix is designed based on the location or spatial attributes of the household from the hospital closest to it. Using the geocoded household addresses and hospital locations in the counties, three distance buffers (1 mile apart) are designed around the hospitals to measure the potential benefits associated with a household being closer to a hospital. Households that fall within the buffer zones are assigned an adjacency value of one (1); all other households are assigned adjacency

values of zero (0). This process yields weight matrixes[5] that are used to define three spatial dummy variables: SPATIAL1, if a household falls within the one-mile distance buffer around the hospital; SPATIAL2, if a household falls within the two-mile buffer; and SPATIAL3, if the household falls within the three-mile distance buffer around the hospital. The reference category is the one-mile distance buffer (SPATIAL1)[6].

Turning to the measurement of life satisfaction, many studies, particularly those conducted by psychologists, have used additive indexes of some kind to measure satisfaction (Ellen and Turner, 1997; Vanfossen, 1981; Gove et al., 1983; Sousa-Poza and Sousa-Poza, 2000; Lu, 1999). In compiling these indexes, researchers first determine a list of personal or neighborhood attributes that are deemed important to individuals. Then they ask respondents to rate the attributes on a Likert-type scale, i.e., to express the extent of their agreement or disagreement with statements that reflect positive or negative attitudes toward these attributes. The ratings are then added up to generate an aggregate measure (Clark and Oswald, 1994). Such aggregate measures of satisfaction have been deemed unreliable (Golant, 1982; Gerdtham and Johannesson, 1997) because expressed satisfaction represents integrated participant perceptions that embrace a wide range of external conditions.

Moore (1986) posits that a reacting individual is likely to attach different levels of importance to various attributes of his/her life and their weights are not likely to be well understood. Therefore, it becomes difficult if not impossible to construct externally calculated reliable measures of life satisfaction. To avoid the above-mentioned

[5] The first-order contiguity matrix is used (row-standardized so that each row's elements sum to one). The prestandardized form of this matrix, WP is defined such that wp_{ij}, is 1 if the i^{th} and the j^{th} observation share a common buffer, and 0 otherwise.

[6] See Bukenya (2001) for a detailed discussion of the spatial weight matrix.

complication, quality of life satisfaction and health, the dependent variables in this paper, are measured using a single question as described below.

Quality of life satisfaction measurements are obtained by asking a categorical question and rating the responses on a three-point scale (Table 1). The ranking of the responses provides three levels of satisfaction. Since the full comparability of utility functions is assumed, the satisfaction derived from the three utility levels is taken to be the same for all individuals. Following the same approach, the health index is also represented by a categorical measure as shown in Table 1. This type of categorical health measure has been shown to capture important information about the individual's health (Connelly et al. 1989) and to be an important predictor of mortality (Wannamethee and Shaper, 1991; Kaplan and Camacho, 1983; Idler and Kasl, 1991). The health index variable is also used as an independent variable in the estimation of the structural quality of life equation, by entering two dummy variables for fair health and good health.

Empirical Results

The empirical results suggest that the models explain a substantive amount of the variations in the dependent variables. The goodness of fit value (pseudo R-square) is 0.299 and 0.318 in the quality of life and health structural equations, respectively and 0.295 for the reduced form quality of life model. It is imperative to note, however, that the pseudo R-square as a measure of goodness of fit deserves only limited attention, because it is chosen to maximize the joint density of the observed dependent variables rather than maximizing a criterion based on prediction of y, as with R^2 in OLS regression analysis. To further examine the goodness of fit of the estimated ordered probit equations, frequencies of actual and predicted outcomes are reported in Table 2.

Table 1: Summary of Variables

1. Dependent variables

a.) CATEGORICAL QUALITY OF LIFE SATISFACTION Assessment of personal satisfaction:
0 = the daily life is never a source of personal satisfaction,
1 = the daily life is sometimes a source of personal satisfaction,
2 = the daily life is a source of personal satisfaction most of the time.

b.) CATEGORICAL HEALTH INDEX Assessment of own health on a three-point scale:
0 = poor health, 1 = fair health, 2 = good health.

2. Independent variables

MALE	= 1 if male
RACE	= 1 if white
AGE1	= 1 if age is 18-34 years
AGE2	= 1 if age is 35-44 years
AGE3	= 1 if age is 45-64 years
AGE4	= 1 if age is >64 years
UNEMPLOYM	= 1 if unemployed
CIVIL STATUS	= 1 if the individual is not married or cohabiting
HPROBLEMS	= 1 if the parents or siblings had any health problems
EDUC1	= 1 if less than high school education
EDUC2	= 1 if high school education
EDUC3	= 1 if university education
INC1	=1 if the gross annual income is in the first quartile of the income distribution, i.e., < $20,000
INC2	= 1 if the gross annual income is in the second quartile of the income distribution, i.e., $20,000-$25,353
INC3	= 1 if the gross annual income is in the third quartile of the income distribution, i.e.,$ 25,354-$34,075
INC4	= 1 if the gross annual income is in the fourth quartile of the income distribution, i.e., > $34,076
RELIGION	= 1 if religion is said to be a source of strength and comfort
AMENITIES	= 1 if there is a hospital or college/University in the county of residence
HEALTH0	= 1 if the health status is rated as bad health in the categorical health question
HEALTH1	= 1 if the health status is rated as fair in the categorical health question.
HEALTH2	= 1 if the health status is rated as good in the categorical health question
LOCAL GOVT.	= 1 if the performance of the county government is rated as fair or good
REGION	= 1 if the individual lives in a county located in Southern WV.
GOVT. EDUC. EXP	= Government expenditure on education and training assistance in 1998
YEARS OF RESIDE	= 1 if more than 10 years of residence
NEIGHBORHOOD	= 1 if satisfied with the neighborhood as a place to live
ENVIRONMENT	= 1 if concerned with the quality of the environment in the county
SPATIAL EFFECT	= 1 if the household falls within the designed distance buffers

The results in Table 2 suggest that the structural and reduced form equations correctly predict 683 (66 percent) and 684 (67 percent), respectively, of the 1028 observations. The structural form equation predicts that 429 (observed: 383) of the total 1028 respondents fall in the higher satisfaction category. The reduced form equation predicts that 477 (observed: 383) of the total respondents fall in the highest satisfaction category. In general, however, the two models perform similarly with the structural model correctly predicting 66% of the outcomes and the reduced model predicting 67%.

Table 2: Frequencies of Actual and Predicted Outcomes for OrdProbt Model

Structural Form Quality of Life Equation

	Predicted			
Actual	0	1	2	Total
0	346	4	1	351
1	7	98	189	294
2	23	121	239	383
Total	376	223	429	1028
Model Prediction				66%

Reduced Form Quality of Life Equation

	Predicted			
Actual	0	1	2	Total
0	349	1	1	351
1	31	61	202	294
2	66	43	274	383
Total	446	105	477	1028
Model Prediction				67%

The predicted percentages are calculated as: (predicted/total sample)*100

To control for heteroscedasticity in the data, the variance of the error terms is assumed to take the form: $\mathrm{Var}[\varepsilon_i] = [\exp(\gamma' z_i)]^2$. The variables included in Z_i are income and government expenditure. The estimates of the variance functions based on the above specification suggest that income and government expenditure significantly explain the variation in the disturbance variances across observations. The estimated effects of each independent variable on the dependent variable are discussed and summarized in Tables 3 and 4. To facilitate the interpretation of the results, the predicted probabilities of being in

the highest quality of life satisfaction category for each level of the explanatory variable are also reported in Table 5 at the mean level of all explanatory variables.

Looking at the effects of the demographic variables, age, gender, and marital status appears to be associated with both health and quality of life satisfaction. On the other hand, though race appear as significant factor in influencing an individual's satisfaction or dissatisfaction with life, it does not appear to have a significant effect on a person's health status in the structural models. The race-health results are somewhat surprising. Based on the general US statistics, race (particularly being Black) is negatively correlated with health. However, our findings show a positive relationship. The possible explanation might lie in the fact that the population of West Virginia is overwhelmingly white, such that the effect of the black race is not readily captured.

The coefficient of the age (AGE4) variable is negative and statistically significant in both equations. This implies that older individuals are less satisfied with their quality of life, and are more likely to be in poor health status than individuals in the youngest age group (18 to 34 years). The impact of age on health is, of course, expected, but the negative effect of age on satisfaction with quality of life, is not equally obvious. The explanation is probably to be found in the observation that the health status of individuals in the higher age group (AGE4) largely influences quality of life satisfaction. When health status is assumed constant, the sign of the AGE4 coefficient becomes positive, implying that when controlled for health status, individuals in the highest age group are more likely to be satisfied with life compared to individuals in the youngest age group.

The effect of AGE2 (35 to 44 years) and AGE3 (45 to 64 years) is positive, but only the coefficient on AGE3 is significant in the quality of life equation. The overall results

Table 3: Ordered Probit Maximum Likelihood Estimation: Structural Form Equations

	QOL Equation		Health Equation	
Variables	Coefficient	t-ratio	Coefficient	t-ratio
ONE	1.984***	4.509	2.016***	3.775
MALE	-0.530	-1.515	-0.281*	-1.645
SINGLE	-0.572*	-1.703	-1.233*	-1.883
UNEMPLOYM	-0.280***	-4.043	-0.285***	-2.892
RELIGION	1.345*	1.765	0.715**	1.890
AGE2	0.058	1.245	0.457e-01	0.928
AGE3	0.124*	1.761	-0.185e-02	-0.658
AGE4	-0.587*	-1.722	-1.089***	-2.550
EDUC2	0.217**	2.250	0.198**	2.301
EDUC3	1.171*	1.668	0.49e-02 *	1.658
WHITE	1.496**	2.102	0.026	0.149
INC2	0.284**	2.056	1.643**	2.095
INC3	0.175**	2.429	0.186***	2.533
INC4	0.750***	6.408	0.226**	1.932
HEALTH1	0.197***	-2.947	-------	-------
HEALTH2	0.269***	3.824	-------	-------
HPROBLEMS	------	------	-0.211***	-2.970
DURATION2	-0.027	-1.170	------	------
LOCAL GOVT. SERVICES	0.301	1.309	0.142	1.157
ENVIRON. SERVICES	-0.003	-1.399	-0.516*	-1.746
REGION	0.033	0.359	-0.265*	-1.633
NEIGHBORHOOD	1.334*	1.715	0.538*	1.602
AMENITIES	1.171	1.638	0.608e-04**	2.053
SPATIAL2	0.069	0.739	1.345*	1.757
SPATIAL3	-0.046	-0.598	-0.396***	-2.357
GOVT. EXP. ON EDUC.	0.016	0.279	0.181e-01	0.991
μ_1	1.011***	19.459	1.294***	14.335
Interactions completed	21		22	
Sample size	1028		1028	
Log-L	-954.8		-961	
Model χ^2	26.02		12.50	
Pseudo R^2	0.299		0.318	
DF	1003		1005	

*** p<.01, ** p<.05, * p<.10.

Table 4: Ordered Probit Maximum Likelihood Estimation: Reduced Form Equation -- Dependent variable: QOL Satisfaction Index

Variables	Coefficient	t-ratio	P-value
ONE	2.730***	3.823	0.0001
MALE	-0.276***	-3.149	0.0016
SINGLE	-2.167**	-2.165	03046
UNEMPLOYM	-2.099***	-3.580	0.004
RELIGION	0.619**	1.924	0.0544
AGE2	-0.307	-1.341	0.1799
AGE3	-0.0001**	-1.895	0.0581
AGE4	0.003	1.383	0.1666
EDUC2	0.334***	3.162	0.0017
EDUC3	1.786***	10.537	0.0000
WHITE	7.959e-006	0.345	0.7303
INC2	0.308***	3.291	0.0010
INC3	0.267***	3.176	0.0015
INC4	0.005***	2.435	0.0149
HPROBLEMS	-0.769***	-6.640	3.14e-011
DURATION2	0.062	1.318	0.1875
LOCAL GOVT SERV.	0.124	1.504	0.1326
ENVIRONMENT SERV.	-0.658**	-2.024	0.0429
REGION	-0.124*	-1.616	0.1062
NEIGHBORHOOD	0.578*	1.698	0.0895
AMENITIES	1.349*	1.775	0.0759
SPATIAL2	0.058	0.905	0.3656
SPATIAL3	-0.587*	-1.746	0.0808
GOVT. EXPENDITURE	1.132	1.612	0.1071
μ_1	4.762***	5.259	1.45e-007

Interactions completed	23
Sample size	1028
Log-L	-954
Model χ^2	13.7640
Pseudo R^2	0.295
DF	1004

*** p<.01, ** p<.05, * p<.10.

Table 5: Predicted Probabilities

Variable	Direct Effects	Total Effects
MALE	0.51230	0.55211
FEMALE	0.54872	0.60145
SINGLE	0.47199	0.01882
MARRIED	0.62010	0.73251
EMPLOYED	0.52343	0.58647
UNEMPLOYED	0.43550	0.49124
RELIGIOUS	0.53650	0.55111
NOT RELIGIOUS	0.46219	0.48540
AGE1	0.59178	0.55821
AGE2	0.57284	0.63417
AGE3	0.55390	0.96011
AGE4	0.60309	0.63010
EDUC1	0.57221	0.55415
EDUC2	0.51300	0.62832
EDUC3	0.61071	0.67180
WHITE	0.57960	0.57260
NON WHITE	0.45607	0.49113
INC1	0.52084	0.53199
INC2	0.58016	0.57879
INC3	0.61884	0.58871
INC4	0.59874	0.61343
HPROBLEMS	0.38112	0.36512
NO HPROBLEMS	0.61414	0.52823
HEALTH0	0.50000	0.38365
HEALTH1	0.51192	0.44145
HEALTH2	0.59587	0.62234
DURATION1	0.55121	0.52211
DURATION2	0.55079	0.54140
LOCAL GOVT. SERVICES=0	0.38100	0.39430
LOCAL GOVT. SERVICESS	0.54675	0.59390
ENVIRONMENT SERVICES=0	0.38767	0.32861
ENVIRONMENT SERVICES	0.56501	0.48741
SWV REGION	0.54199	0.39784
EPH REGION	0.55590	0.41541
NEIGHBORHOOD=0	0.44133	0.46933
NEIGHBORHOOD	0.52084	0.53315
AMENITIES=0	0.04836	0.50515
AMENITIES	0.50107	0.55786
SPATIAL1	0.54646	0.54650
SPATIAL2	0.41234	0.43711
SPATIAL3	0.28987	0.38544
GOVORNMENT EXPENDITURE	0.41317	0.49329

Predicted Probabilities of daily life being a source of personal satisfaction most of the time
The predicted probabilities are calculated as $F(\hat{\beta}'X) = \hat{F}$ (Green, 1993: p.645).

suggest a U-shaped relationship between age and quality of life satisfaction, when health status is controlled for. The predicted probability of being satisfied most of the time in the structural equation is 0.59 in the age-group 18 to 34 years, 0.57 in the age-group 35 to 44 years, 0.55 in the 45 to 64 years and 0.60 in the age group over 65 years.

The results of the marital status variable suggest that being single has a negative effect on both health status and quality of life satisfaction. These results conform to the findings of Gove et al., (1993) who found that a strong causal relationship between being married and quality of life. The total effect in the reduced form equation is also negative and statistically significant because of the direct and indirect effects manifested in the structural form equations. The predicted probability of being satisfied most of the time is 0.47 for persons who are single and 0.62 for persons who are married or cohabiting. The effect on satisfaction of being single is greater compared to the difference in satisfaction between the highest and the lowest income quartile ($P<0.05$) and the difference in satisfaction between men and women ($P<0.01$). However, the effect does not differ significantly from the difference in satisfaction between the highest and the lowest education category.

Male gender and white race have a negative and positive effect, respectively, in both quality of life and health equations, though only white race is statistically significant in the structural equation. In the reduced form equation, both variables maintain their signs, but show a reverse in significance. The predicted probability of being satisfied most of the time is 0.55 for men and 0.60 for women whereas it is 0.57 for being of white race and 0.49 for a non-white race in the reduced form quality of life equation.

The estimated coefficients of the income variables (INC2, INC3, and INC4) are positive and statistically significant in all equations. As predicted by economic theory, the results suggest that individuals with higher incomes are more likely to be satisfied with life and to have better health, other things being equal. The predicted probability of being satisfied most of the time increases from 0.53 in the lowest income quartile to 0.61 in the highest income quartile, taking into account the total effect of income. If only the direct effect of income on satisfaction is taken into account, the difference in the predicted probability for a respondent being satisfied most of the time is only two percentage units between the lowest and the highest income quartiles. This indicates that a large part of the effect of income on quality of life satisfaction occurs through the intervening health variable (Bezruchka, 2001)—probably through health expenditures such as health insurance and drug costs (Nixon, 1997).

The results for the education variable are positive, as hypothesized, and highly significant in both the structural and reduced form model. The results are compatible with theory and with societal expectations that higher education attainment is associated with improved socioeconomic status, higher wage rates, better health, all of which lead to better living standards. The predicted probability of being satisfied most of the time increases from 0.55 with less than high school education to 0.67 for university education, based on the total effect in the reduced form model. If only the direct effect on quality of life satisfaction is considered, the probability of being satisfied most of the time would increase from 0.57 (less than high school education) to 0.61 (university education). The difference in satisfaction between having a university education and having less than high school education is greater than the difference in satisfaction between men and women,

but not significantly different from the difference in satisfaction between the highest and the lowest income quartile.

Unemployment, as previously hypothesized, has a negative and statistically significant correlation with health status and quality of life satisfaction. Similar to income and education, unemployment affects health status and quality of life satisfaction simultaneously. Labor income is the primary source of income for the majority of households. It is determined by the unemployment rates, number of hours worked, labor participation rate, etc. Thus, unemployment results in lower levels of quality of life and health care. It must be noted however, that one could make an argument for reverse causation, i.e., work status (and unemployment) being partially the result of health problems. The predicted probability of being satisfied most of the time is 0.49 for individuals who are unemployed and 0.58 for employed individuals. The effect of unemployment on satisfaction is not significantly different from the effect of gender or the effect of being single. The effect is also not significantly different from the difference in satisfaction between the highest and the lowest income quartile.

Health status has a significant and positive effect on quality of life satisfaction. Both health dummy variables are significant and the effect of good health (HEALTH2) is significantly higher than the effect of fair health (HEALTH1). The predicted probability of being satisfied most of the time is 0.38 with a bad health status and 0.62 with a good health status. In the reduced form model, the results suggest that improving a person's health status from fair to good health, for instance, would increase the probability of being "satisfied with life most of the time" by 0.18. The proxy variable for inherited health status, health problems in the family (HPROBLEMS), is significant and negatively

related to health status in the structural form equation. In the reduced form equation, initial inherited health has the hypothesized negative relationship on quality of life satisfaction, but not statistically significant. The predicted probability of being satisfied most of the time in the reduced form equation is 0.36 if the parents or siblings had any health problems (HPROBLEMS) and 0.53 in the absence of any health problems (NO HPROBLEMS).

The variable for religion shows a positive impact in the structural model, but the coefficients are significant only in the health equation. In the reduced form model, the total effect of religion on quality of life satisfaction is positive and significant. Historically, religion has been recognized as a powerful factor in promoting good health among individuals (Yinger, 1957) and the whole society (Durkheim, 1976). This presumption has endured in the theoretical literature, at least in part, because, as Ellison (1991) and Idler (1987) have suggested, religion appears to provide a variety of inducements to personal and community well-being, such as enhanced social integration and support. Idler (1987) adds that involvement in religion might also constrain high-risk behavior, such as smoking, drinking and sexual activity outside a stable relationship.

However, not all studies have supported these hypotheses about the religion-health link. A few studies have shown no relationship between the two, especially when such factors as social class and previous health status are controlled for. Some scholars (Singer, 1979) have even argued that religiosity may have deleterious effects on overall health. Alternatively, individuals in poor health may be drawn to religion in an effort to seek comfort and healing, raising the question of causality. Thus, the overall contribution of religion to well-being remains a source of controversy. The predicted probability of

being satisfied most of the time by someone who is religious is 0.55 in the reduced form equation, while the probability of being satisfied most of the time by someone who is not religious is 0.48.

The duration variable (i.e., years of residence) is included to examine the effect that living in the county for more than ten years might have on satisfaction. The estimated results do not have the hypothesized positive effect in the structural model and are not statistically insignificant. In the reduced form model, the variable has the hypothesized positive sign though not statistically significant. The influence of the intervening health variable is a possible explanation for the positive sign in the reduced form equation.

Our analysis of the policy variables show that government expenditure on education and local government services has a positive, but not significant, effect on health and quality of life satisfaction. The results imply, for instance, that other things being equal, government expenditure on education and training would increase people's satisfaction. Similarly, the local government services variable is not statistically significant. The predicted probability of being satisfied most of the time in the reduced form equation is 0.59 with good local government services and 0.39 with poor local government services.

As hypothesized, the amenity variable has a positive and significant effect in the structural form quality of life model. The amenity variable has a highly significant positive effect in the health equation, suggesting that the presence of amenities in a region can improve one's health status. In the reduced form equation, the intervening positive effect through the health equation increases the total effect of amenities on quality of life satisfaction. The predicted probability of being satisfied most of the time is

0.55 in counties with medical facilities, higher institutions of learning, or closer to metropolitan areas, and 0.51 in counties without such facilities. In the literature, amenities are closely related to the environment, but our results for the environmental variable are contrary to the hypothesized positive sign.

An analysis of the geographical variables indicates that there is evidence to suggest that location and distance play a role in determining whether an individual will report better health status or not. Given the nature of our data, three spatial variables are included to control for spatial dependence that might influence the analysis. The results suggest that the effect of the spatial variable is more pronounced in the health equation. In that equation, the SPATIAL2 variable has a positive and significant effect, and SPATIAL3 has a highly significant negative effect. This suggests that individuals residing at a distance beyond a three-mile radius from a hospital are more likely to report poor health status than people living within a one-mile radius. With regard to quality of life, the results in the reduced form equation show a negative effect for SPATIAL3. The predicted probability of being satisfied most of the time in the reduced form equation decreases from 0.54 with SPATIAL1 to 0.43 and to 0.38 for SPATIAL2 and SPATIAL3, respectively.

An analysis of the region variable indicates that there is weak evidence to suggest that individuals living in the southern region are more likely to be satisfied with life than those living in the Eastern Panhandle region. Although the coefficient for this variable is positive in the structural quality of life model, it is not statistically significant. The coefficient of the regional variable is negative and statistically significant in both the health and reduced form quality of life model. The predicted probability of being satisfied most of the time is 0.39 for living in the southern region and 0.41 for living in the eastern

panhandle. Overall the results suggest that individuals living in the southern region are less likely to be satisfied with life compared to those residing in eastern panhandle.

Lastly, neighborhood appears to play a significant role in enhancing both quality of life satisfaction and health status. In the structural form model, the coefficient for the neighborhood variable is positive and statistically significant. In the reduced form model, the neighborhood variable maintains both a positive sign and the significance level. The predicted probability of being satisfied most of the time is 0.53 with good neighborhood and 0.46 with bad neighborhood.

Summary and Conclusions

It is interesting to briefly relate our results to two recent studies (Gerdtham and Johannesson, 1997; Clark and Oswald 1994) that used a similar methodology. Looking at the similarities, the results in our study conform to the results in these two studies in several areas. First, the two studies also found that unemployment and being single have strong negative effects on quality of life satisfaction (happiness). Second, they also found a U-shaped relationship between age and quality of life satisfaction, with quality of life satisfaction/happiness being lowest for individuals in their mid-thirties (Clark and Oswald, 1994) and in the age-group 45 to 64 years in our study and in Gerdtham and Johannesson (1997).

There are, however, some important differences in the results. For instance, using Swedish macro data, Clark and Oswald (1994) found no systematic relationship between income and quality of life satisfaction/happiness, and they found a negative relationship between education and quality of life satisfaction. As predicted by economic theory, our study (similar to Gerdtham and Johannesson, 1997) found a positive relationship between

quality of life and both income and education. Clark and Oswald (1994) also found that men were happier than women, whereas in our study and in Gerdtham and Johannesson (1997), the reverse is observed.

A component that differentiates our study from these two previous studies is that we try to control for spatial effects in our analysis through the inclusion of spatial variables. We acknowledge that testing and correcting for spatial dependence in discrete choice models is technically more demanding than addressed in this paper (see Smith and LeSage, 2001; Anselin, 2001; Bell and Bockstael, 2000; Fleming, 2002) such that we can not fully attribute the observed differences in the results on correction of spatial effects. However, by including spatial variables in our analysis we: (1) Allow greater insights into the nature and accuracy of the examined data. (2) Provide a more detailed understanding of the nature of relationships and their variation over space. (3) Demonstrate the possible naiveté of conventional approaches to data analysis that often ignores spatial non-stationarity (Fortheringham and Wegener, 2000: 25-26).

In sum, the results in this study support the argument that many ”non-economic” variables are as important for quality of life satisfaction as income and consumption. The results suggest that socioeconomic variables such as unemployment, health status, gender, marital status, regional differences and education are as important as income in determining one's satisfaction with the life they lead. These findings have important implications (for instance, in studies of equity and distributional issues) since it is important to also take into account the distribution of non-economic factors such as education, health status and employment possibilities when assessing the distribution of welfare in society.

REFERENCES

Amemiya, Takeshi. 1981. "Qualitative response model: a survey." *Journal of Economic Literature* 19, 481--536.

Anselin, L. 2001. "Spatial econometrics," in B. Baltagl (ed.), *A Companion to Theoretical Econometrics*, Oxford: Basil Blackwell, pp. 310-330.

Anselin, L., Griffith, D.A., 1988. "Do spatial effects really matter in regressional analysis?" *Papers of the Regional Science Association* 65, 11--34.

Anselin, L., 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, Kluwer.

BBER. 2000. "County data profile." Bureau of Business and Economic Research, West Virginia University. http://www.be.wvu.edu/serve/bureau/online/cdp.htm

Bell, K.P., Bockstael, N.E., 2000. "Applying the generalized moments estimation approach to spatial problems involving microlevel data." *The Review of Economics and Statistics* 82, 1: 72--82.

Bezruchka, Stephen (2001). "Health and income equity," International Health Program University of Washington and Health Alliance International. http://depts.washington.edu/eqhlth/index.html.

Boadway, R.W., Bruce, N., 1984. *Welfare Economics*. Blackwell, Oxford.

Bukenya, J.O., 2001. "An analysis of quality of life, income distribution and rural development in West Virginia." *Dissertation*: Division of Resource Management, West Virginia University. Morgantown.

Cameron, C., Trivedi, P., 1986. "Econometric models based on count data: comparisons and applications of some estimators and tests," *Journal of Econometrics* 1, 29--53.

Clark, A.E., Oswald, A.J., 1994. "Unhappiness and unemployment." *Economic Journal* 104, 648--659.

Cliff, A. D., Ord, J.K., 1973. *Spatial Autocorrelation*. London, Pion.

Connelly, J.E., Philbrick, J.T., Smith, R., Kaiser, D.L., Wymer, A., 1989. "Health perceptions of primary care patients and the influence on health care utilization." *Supplement to Medical Care* 27, 99--109.

Dissart, J.C., Deller, S.C., 2000. "Quality of life in the planning literature." *Journal of Planning Literature* 15, 1: 135--61.

Doreian, Patrick. 1980. "Linear models with spatially distributed data: Spatial disturbances or spatial effects?" *Sociological Methods and Research* 9, 29--60.

Doreian, Patrick. 1981. "Estimating linear models with spatially distributed data." In: Leinhardt, S., (Ed.), *Sociological Methodology*. San Francisco: Jossey-Boss, 359--88.

Durkheim, Emile. 1976. *The Elementary Forms of Religious Life*. London, George Allen and Unwin.

Dustman, C., 1996. "The social assimilation of immigrants." *Journal of Population Economics* 9, 37--54.

Easterly, William. 1999. "Life during growth." World Bank. http:www.worlbank.org/html/prdmg/grthweb/growth_t.htm.

Ellen, I.G., Austin, M.T., 1997. "Does neighborhood matter? Assessing recent evidence." *Housing Policy Debate* 8, 4: 833--66.

Ellison, C.G., 1991. "Religious involvement and subjective well-being." *Journal of Health and Social Behavior* 32, 80--99.

ESRI ArcData Online. 2000. "GIS and Mapping software." GIS data on the web. http://www.esri.com/data/online/index.html.

Fleming, M. 2002. "A review of the techniques for estimating spatially dependent discrete choice models." In Anselin and Florax (eds.), *Advances in Spatial Econometrics*, Forthcoming.

Fortheringham, A.S., Wegener, M., 2000. *Spatial Models and GIS: New Potential New Methods*. London: Taylor and Francis.

Gerdtham, U., Johannesson, M., 1997. "The relationship between happiness, health and socio-economic factors: Results based on Swedish micro data." Economic and Finance Working Paper No. 207. Stockholm School of Economics. Stockholm Sweden.

Golant, S.M., 1982. "Individual Differences underlying the dwelling satisfaction of the elderly." *Journal of Social Issues* 38, 121--133.

Gove, W.R., Hughes, M., Style, C.B., 1983. "Does marriage have positive effects on the well-being of the individual?" *Journal of Health and Social Behavior* 24, 122--131.

Greene, H.W., 1993. *Econometric Analysis*, Second Edition, Macmillan Publishing Company, New York.

Greene, H.W., 1995. *LIMDEP TM Version 7.0 User's manual*. Econometric Software Inc.

Hall, E.R., Jones, C.I., 1997. "Levels of economic activity across Countries." *American Economic Review* 87, 2: 173--177.

Hall, E.R., Jones, C.I., 1998. "Why do some countries produce so much more output than others?" Stanford University, mimeo.

Hanushek, E.A., Jackson, J.E., 1977. *Statistical Methods for Social Scientists*. New York: Academic Press.

Heubusch, Kevin. 1998. "Small is beautiful." *American Demographics* 20, 1: 43--49.

Idler, E.L., 1987. "Religious involvement and the health of the elderly: Some hypotheses and an initial test." *Social Forces* 66, 226--238.

Idler, E.L., Kasl, S., 1991. "Health perceptions and survival: Do global evaluations of health status really predict mortality?" *Journal of Gerontology* 46, 555--565.

Kaplan, G.A., Camacho, T., 1983. "Perceived health and mortality: A nine-year follow-up of the human population laboratory cohort." *American Journal of Epidemiology* 117, 292--304.

Killian, M., Tolbert, C.Jr., 1993. "Mapping social and economic space." In: Singelmann, J., Deseran, F.A., (Eds.), *Inequalities in Labor Market Areas*. Boulder CO: Westview Press, pp. 69--82.

Land, K., Deane, G., 1992. "On the large-sample estimation of regression models with spatial or network effects terms: A two-stage least squares approach." In: Marsden, P., (Ed.), *Sociological Methodology*, Cambridge MA, Blackwell, pp. 221--248.

LeSage, James. 1999. "Spatial Econometrics." In: Loveridge, S., (Ed.), *The Web Book of Regional Science*, Regional Research Institute, West Virginia University. http://www.rri.wvu.edu/regscbooks.htm

Lu, M., 1999. "Determinants of residential satisfaction: Ordered logit vs. regressional models." *Growth and Change* 30, 2: 264--84.

Lyson, A.T., Tolbert, C.Jr., 1996. "Small manufacturing and civic welfare in the U.S. nonmetropolitan counties: A regional comparison." *Environment and Planning* 28, 1779--1794.

Mencken, F.C., 1997. "A study of the region: Regional differences in socioeconomic well-being in Appalachia." *Sociological Focus* 30, 79--97.

Mencken, F.C., 1998. "Persistent differences in well-being between Appalachia subregions." *Growth and Change* 29, 469--479.

Moore, E., 1986. "Mobility intention and subsequent relocation." *Urban Geography* 7, 497--514.

Nixon, John. 1997. "Convergence analysis of health care expenditure in the EU countries using two approaches." Center for Review and Dissemination and Department of Economics and Related Studies. UK: University of York.

Ord, Keith. 1975. "Estimation methods for models of spatial interaction." *Journal of the American Statistical Association* 70, 120--26.

Reichert, C.V., Rudzitis, G., 1992. "Multinomial logistic models explaining income changes of migrants to high-amenity counties." *Review of Regional Studies* 22, 1: 25--42.

REIS. 1969-1998. "State and county level variables." Regional Economic Information System. http://fisher.lib.Virginia.EDU/reis/

Schuessler, K.F., Fisher, G.A., 1985. "Quality of life research and sociology." *Annual Reviews in Sociology* 11, 129--49.

Singer, M.T., 1979. Coming out of the cults. *Psychology Today* (January), 72--82.

Smith, T. E., and J. P. LaSage. 2001. A bayesian Probit Model with spatial dependences. Manuscript available at http:///www.spatial-econometrics.com.

Sousa-Poza, A., Sousa-Poza, A.A., 2000. "Well-being at work: A cross-national analysis of the levels and determinants of job satisfaction." *Journal of Socio-Economics* 29, 517--538.

Vanfossen, B.E., 1981. "Sex difference in the mental health effects of spouse support and equity." *Journal of Health and Social Behavior* 22, 130--143.

Wannamethee, G., Shaper, A.G., 1991. "Self-assessment of health status and mortality in middle-aged British men." *International Journal of Epidemiology* 20, 239--245.

Whittle, P., 1954. "On stationary processes in the plane." *Biometrika* 41, 434--449.

Yinger. J.M., 1957. *Religion, Society, and the Individual*. Macmillan, New York.