2006

Modeling Small Business Growth, Migration Behavior, and Household Income in Appalachia: A Spatial Simultaneous Equations Approach

Gebremeskel H. Gebremariam

Tesfa Gebremedhin
tgebreme@wvu.edu

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

Modeling Small Business Growth, Migration Behavior, and Household Income in Appalachia: A Spatial Simultaneous Equations Approach

By

Gebremeskel H. Gebremariam[1]
Tesfa G. Gebremedhin
Peter V. Schaeffer

RESEARCH PAPER 2006-8

Abstract: In this paper, a spatial simultaneous equations model of business growth, migration behavior and median household income is empirically estimated. The empirical simultaneous model is developed from the equilibrium relationships among these variables where each variable is assumed to adjust to its equilibrium level with a substantial lag through a partial equilibrium adjustment process. We use Generalized Spatial Three-Stage Least Squares estimator to estimate the empirical model using data from 418 Appalachian counties for 1990-2000. The results suggest the existence of very strong interdependences among business growth, migration behavior and median household income in the form of *feed-back simultaneities*, *spatial autoregressive lag* simultaneities and *spatial cross-regressive lag* simultaneities.

Key Words: Appalachia, Simultaneous Model, Small Business, Migration, Household Income

[1] Graduate Research Assistant, Professor, and Professor/Director, Division of Resource Management, Davis College of Agriculture, Forestry & Consumer Sciences, West Virginia University.

Selected Paper presented at the Northeastern Agricultural and Resource Economics Association Annual Meeting, Mystic, Connecticut, June 11-14, 2006. We acknowledge the review comments of Dale Colyer.

This research was supported by Hatch funds appropriated to the Agricultural and Forestry Experiment Station, Davis College of Agriculture, Forestry and Consumer Sciences, West Virginia University.

Modeling Small Business Growth, Migration Behavior, and Household Income in Appalachia: A Spatial Simultaneous Equations Approach

Introduction

Persistent poverty is one of the most critical social problems facing policy makers in the United States. Despite decades of government intervention and the spending of billions of public funds, many communities still remain in poverty. The economic boom of the 1990s failed to reduce poverty in many counties (Rupasingha and Goetz, 2003) and counties in the Appalachia had above average poverty rates in 1990s. In spite of an expanding economy, many counties in Appalachia suffer from high unemployment rates, shrinking economic bases, deeply rooted poverty, low human capital formation, and out-migration (Deavers and Hoppe, 1992; Hayness, 1997; Dilger and Witt, 1994**;** Maggard, 1990). The slow growth of income and employment, out-migration and disappearance of rural households are both causes and effects of persistent high rates of poverty. Lagging economic development negatively affects the economic and social well-being of rural populations, the health of local businesses, and the ability of local governments to provide basic human services (Cushing and Rogers, 1996).

The changing structure of traditional industries and the impact of those changes on local communities are sources of concern to many groups interested in the welfare of rural areas. State policy makers and local leaders place a high priority on local economic development (Pulver, 1989; Ekstrom and Leistritz, 1988). Consequently, a better understanding of factors that influence local employment, earning capacity and quality of life issues has become important from county, state and regional policy perspectives in designing human capital development programs for rural community development. Since many of the forces responsible for past economic and social changes in rural communities will continue to affect rural families, it is important to study the rural economy and evaluate alternative policy measures to promote diverse and resilient local communities.

Improving the economic base of a region requires an economic environment where business can prosper. Appalachia, despite efforts of multilateral, national and local policy programs to induce economic prosperity and ameliorate poverty, has many economically depressed communities. To strengthen and diversify their economies, policy makers and local leaders need to know the characteristics and impacts of small businesses on the local economy. Understanding the characteristics of poverty and the contribution of small businesses to the growth of local economies is crucial for designing specific development policies.

Studies have shown that employment, occupation and salary as well as the quality and quantity of local public services are major considerations in the decision to migrate (Clark and Hunter, 1992; White and Knapp, 1994). The direction and magnitude of migration depends on salary differentials across regions. High earnings are associated with net in-migration and low earnings with net out-migration. Since low-income states are dominated by occupations with relatively low earnings, and the earnings within particular occupations in low-income states tend to be lower than the national average, low-income regions face a net out-migration. The opposite relationships characterize the high-income states. By promoting the growth and development of small businesses, low-income regions can, however, reduce and ultimately reverse the net out-migration of skilled labor.

Although understanding the interconnections between growth of small businesses, net migration, and the incidence of poverty has been an interest of many researchers, only a few attempts have been made to explain their interdependence in a spatial simultaneous equations framework. In this study, we develop a four-equation spatial simultaneous-equations model with small business growth (using employment growth rate as a proxy), gross in-migration rate, gross out-migration rate, and incidence of poverty (using median household income as a proxy) as endogenous variables of the model.

Model Development

The relationship between economic growth and its determinants have been studied extensively. The issue of whether regional development can be associated with population driving employment changes or employment driving population changes (do 'jobs follow people' or 'people follow jobs'?) has, for example, attracted considerable interest among researchers. Empirical work on identification of the direction of causality in this 'jobs follow people' or 'people follow jobs' literature have resulted in the view that empirical models of regional development often reflect the interdependence between household residential choices and firm location choices (Steinnes and Fischer, 1974). To account for this causation and interdependency, Carlino and Mills (1987) suggested and constructed a two-equation simultaneous system with the two partial location equations as its components. This model has subsequently been used by a number of regional science researchers to examine regional economic growth (Boarnet, 1994; Duffy, 1994; Henry et al., 1997; Duffy-Dino, 1998; Barkley et al., 1998; Henry et al., 1999; Edmiston, 2004). More recently, Deller et al., (2001) have expanded the original Carlino-Mills model to capture explicitly the role of income. According to the theoretical base of utility maximization in the traditional migration literature, households migrate to capture higher wages or income. The model expanded by Deller et al., (2001) is three-dimensional (jobs-people-income) and explicitly traces the role of income in the regional growth process. It also explicitly captures the increasing concerns about job quality as measured by income levels those jobs support. There also have been efforts to model the interactions between employment growth and human migration (MacDonald, 1992; Clark and Murphy, 1996), per capita personal income and public expenditures (Duffy-Deno and Eberts, 1991), net migration, employment growth, and average income or earnings in simultaneous-equations methods (Greenwood and Hunt, 1984; Greenwood et al., 1986; and Lewis et al., 2002).

The theoretical base for the interdependencies between employment, migration behavior and income is the idea that households and firms are both mobile and that household location decisions maximize utility while firm location decisions maximize profits. That is, households migrate to capture higher wages or income and firms migrate to be near growing consumer markets. These actions in turn generate income to the regional (local) economy. The location decisions of firms, however, are expected to be influenced not only by population and income (i.e., growing consumer markets) but also by other factors including local business climate, wage rates, tax rates, local public services, and regional location. Firm location decisions are also influenced by the substantial financial incentives that local governments offer in an effort to create jobs, spur income growth, and enhance the economic opportunities of the local population.

Based upon the above mentioned assumptions, the following central hypotheses are constructed for this research:

1. Business growth, migration behavior and median household income growth are interdependent and are jointly determined by county-level variables;
2. Business growth, migration behavior and median household income growth in any county are conditional upon initial conditions of that county; and
3. Business growth, migration behavior and median household income growth in a county are conditional upon business growth, migration behavior and median household income growth in neighboring counties.

To test these hypotheses, a spatial simultaneous equations model of business growth, migration behavior and household median income is used. Following the Carlino and Mills tradition and building on Deller et al. (2001) and Lewis et al. (2002), the basic model is specified as:

$$\left.\begin{cases} INM_{it}^{*} = f_1(OTM_{it}, EMP_{it}, MHY_{it} \,|\mathbf{X}_{it}^{in}) \\ OTM_{it}^{*} = f_2(INM_{it}, EMP_{it}, MHY_{it} \,|\mathbf{X}_{it}^{ot}) \\ EMP_{it}^{*} = f_3(INM_{it}, OTM_{it}, MHY_{it} \,|\mathbf{X}_{it}^{em}) \\ MHY_{it}^{*} = f_4(INM_{it}, OTM_{it}, EMP_{it}, |\mathbf{X}_{it}^{mh}) \end{cases}\right\} \ldots\ldots\ldots(3.1)$$

where INM_{it}^{*}, OTM_{it}^{*}, EMP_{it}^{*}, and MHY_{it}^{*} are respectively equilibrium levels of gross in-migration, gross out-migration, private business employment, and median household income, and i and t indexes represent county and time, respectively. The vectors of additional exogenous variables included in the respective simultaneous equations system are given by $\mathbf{X}_{it}^{in}$, $\mathbf{X}_{it}^{ot}$, $\mathbf{X}_{it}^{em}$, and $\mathbf{X}_{it}^{mh}$, respectively. Equilibrium levels of gross in-migration, gross out-migration, private business employment and median household income are assumed to be functions of actual levels of the respective right-hand included endogenous variables.

To reduce the effects of the large diversity found in the data used in empirical analysis, a multiplicative (log-linear) form of the model is used. This specification also implies a constant-elasticity form for the equilibrium conditions given in (3.1). A log-linear (i.e., log-log) representation of these equilibrium conditions can thus be expressed as:

$$INM_{it}^{*} = \left(OTM_{it}\right)^{a_1} \times \left(EMP_{it}\right)^{b_1} \times \left(MHY_{it}\right)^{c_1} \times \prod_{k_1=1}^{K_1}\left(\mathrm{X}_{k_1 it}^{in}\right)^{x_{1k_1}}$$

$$\rightarrow \ln\left(INM_{it}^{*}\right) = a_1 \ln\left(OTM_{it}\right) + b_1 \ln\left(EMP_{it}\right) + c_1 \ln\left(MHY_{it}\right) + \sum_{k_1=1}^{K_1} x_{1k_1} \ln\left(\mathrm{X}_{k_1 it}^{in}\right) \quad (4.2a)$$

$$OTM_{it}^{*} = \left(INM_{it}\right)^{a_2} \times \left(EMP_{it}\right)^{b_2} \times \left(MHY_{it}\right)^{c_2} \times \prod_{k_2=1}^{K_2}\left(\mathrm{X}_{k_2 it}^{ot}\right)^{x_{2k_2}}$$

$$\rightarrow \ln\left(OTM_{it}^{*}\right) = a_2 \ln\left(INM_{it}\right) + b_2 \ln\left(EMP_{it}\right) + c_2 \ln\left(MHY_{it}\right) + \sum_{k_2=1}^{K_2} x_{2k_2} \ln\left(\mathrm{X}_{k_2 it}^{in}\right) \quad (4.2b)$$

$$EMP_{it}^{*} = \left(INM_{it}\right)^{a_3} \times \left(OTM_{it}\right)^{b_3} \times \left(MHY_{it}\right)^{c_3} \times \prod_{k_3=1}^{K_3}\left(\mathbf{X}_{k_3 it}^{em}\right)^{x_{3k_3}}$$

$$\rightarrow \ln\left(EMP_{it}^{*}\right) = a_3 \ln\left(INM_{it}\right) + b_3 \ln\left(OTM_{it}\right) + c_3 \ln\left(MHY_{it}\right) + \sum_{k_3=1}^{K_3} x_{3k_3} \ln\left(\mathrm{X}_{k_3 it}^{in}\right) \quad (4.2c)$$

$$MHY_{it}^{*} = \left(INM_{it}\right)^{a_4} \times \left(OTM_{it}\right)^{b_4} \times \left(EMP_{it}\right)^{c_4} \times \prod_{k_4=1}^{K_4}\left(\mathrm{X}_{k_4 it}^{mh}\right)^{x_{4k_4}}$$

$$\rightarrow \ln\left(MHY_{it}\right) = a_4 \ln\left(INM_{it}^{*}\right) + b_4 \ln\left(OTM_{it}\right) + c_4 \ln\left(EMP_{it\ it}\right) + \sum_{k_4=1}^{K_4} x_{4k_4} \ln\left(\mathrm{X}_{k_4 it}^{mh}\right) \quad (4.2d)$$

where a_i, b_i *and* c_i, i=1,...,4 are the exponents on the endogenous variables, x_{ik_j} for $i, j = 1,...,4$ are vectors of exponents on the exogenous variables, $\prod$ is the product operator, and K_i for $i = 1,...,4$ are the number of exogenous variables in the gross in-migration, gross out-migration, employment growth and median household income equations, respectively. The log-linear specification has an advantage of yielding a log-linear reduced form for estimation, where the estimated coefficients represent elasticities. Duffy-Deno (1998) and MacKinnon et al. (1983) also show that, compared to a linear specification, a log-linear specification is more appropriate for models involving population and employment densities.

The literature (Edmiston, 2004; Hamalainen and Bockerman, 2004; Aronsson et al., 2001; Deller et al., 2001; Henry et al., 1999; Duffy-Deno, 1998; Barkley et al., 1998; Henry et al., 1997; Boarnet, 1994; Duffy, 1994, Carlino and Mills, 1987; Mills and Price, 1984) suggests that employment and median household income likely adjust to their equilibrium levels with a substantial lag (i.e., initial conditions). Following the previous literature a distributed lag adjustment is introduced and the corresponding partial-adjustment process for each of the equations given in (3.1) is expressed in the following form:

$$\frac{INM_{it}}{INM_{it-1}} = \left(\frac{INM^*_{it}}{INM_{it-1}}\right)^{\eta_{in}} \rightarrow \ln\left(INM_{it}\right) - \ln\left(INM_{it-1}\right) = \eta_{in}\ln\left(INM^*_{it}\right) - \eta_{in}\ln\left(INM_{it-1}\right) \qquad (4.3a)$$

$$\frac{OTM_{it}}{OTM_{it-1}} = \left(\frac{OTM^*_{it}}{OTM_{it-1}}\right)^{\eta_{ot}} \rightarrow \ln\left(OTM_{it}\right) - \ln\left(OTM_{it-1}\right) = \eta_{ot}\ln\left(OTM^*_{it}\right) - \eta_{ot}\left(OTM_{it-1}\right) \qquad (4.3b)$$

$$\frac{EMP_{it}}{EMP_{it-1}} = \left(\frac{EMP^*_{it}}{EMP_{it-1}}\right)^{\eta_{em}} \rightarrow \ln\left(EMP_{it}\right) - \ln\left(EMP_{it-1}\right) = \eta_{em}\ln\left(EMP^*_{it}\right) - \eta_{em}\left(EMP_{it-1}\right) \qquad (4.3c)$$

$$\frac{MHY_{it}}{MHY_{it-1}} = \left(\frac{MHY^*_{it}}{MHY_{it-1}}\right)^{\eta_{mh}} \rightarrow \ln\left(MHY_{it}\right) - \ln\left(MHY_{it-1}\right) = \eta_{mh}\ln\left(MHY^*_{it}\right) - \eta_{mh}\ln\left(MHY_{it-1}\right) \qquad (4.3d)$$

where the subscript t-1 refers to the indicated variable lagged one period, one decade in this study, and $\eta_{in}, \eta_{ot}, \eta_{em}$ and η_{mh} are the speed of adjustment parameters. They are interpreted as the shares or proportions of the respective equilibrium rate of growth that were realized each period.

Substituting from equations (4.2a) – (4.2d) into equations (4.3a) - (4.3d) gives:

$$\ln\left(INM_{it}\right) - \ln\left(INM_{it-1}\right) = \eta_{in}\left[a_1\ln\left(OTM_{it}\right) + b_1\ln\left(EMP_{it}\right) + c_1\ln\left(MHY_{it}\right) + \sum_{k_1=1}^{K_1} x_{1k_1}\ln\left(\mathrm{X}^{in}_{k_1it}\right)\right] - \eta_{in}\ln\left(INM_{it-1}\right)$$

$$\ln\left(INM_{it}\right) = \beta_{11}\ln\left(OTM_{it}\right) + \beta_{12}\ln\left(EMP_{it}\right) + \beta_{13}\ln\left(MHY_{it}\right) + \sum_{k_1=1}^{K_1}\gamma_{1k_1} l\,n\left(\mathrm{X}^{in}_{k_1it}\right) + (1-\eta_{in})\ln\left(INM_{it-1}\right) \qquad (4.4a)$$

$$\ln\left(OTM_{it}\right) - \ln\left(OTM_{it-1}\right) = \eta_{ot}\left[a_2\ln\left(INM_{it}\right) + b_2\ln\left(EMP_{it}\right) + c_2\ln\left(MHY_{it}\right) + \sum_{k_2=1}^{K_2} x_{2k_2}\ln\left(\mathrm{X}^{ot}_{k_2it}\right)\right] - \eta_{ot}\ln\left(OTM_{it-1}\right)$$

$$\ln\left(OTM_{it}\right) = \beta_{21}\ln\left(OTM_{it}\right) + \beta_{22}\ln\left(EMP_{it}\right) + \beta_{23}\ln\left(MHY_{it}\right) + \sum_{k_2=1}^{K_2}\gamma_{2k_2} l\,n\left(\mathrm{X}^{ot}_{k_2it}\right) + (1-\eta_{ot})\ln\left(OTM_{it-1}\right) \qquad (4.4b)$$

$$\ln\left(EMP_{it}\right)-\ln\left(EMP_{it-1}\right)=\eta_{em}\left[a_3\ln\left(OTM_{it}\right)+b_3\ln\left(INM_{it}\right)+c_3\ln\left(MHY_{it}\right)+\sum_{k_3=1}^{K_3}x_{3k_3}\ln\left(\mathrm{X}_{k_3it}^{em}\right)\right]-\eta_{em}\ln\left(EMP_{it-1}\right)$$

$$\ln\left(EMP_{it}\right)\ =\beta_{31}\ln\left(OTM_{it}\right)+\beta_{32}\ln\left(INM_{it}\right)+\beta_{33}\ln\left(MHY_{it}\right)+\sum_{k_3=1}^{K_3}\gamma_{3k_3}\ln\left(\mathrm{X}_{k_3it}^{em}\right)+(1-\eta_{em})\ln\left(EMP_{it-1}\right) \qquad (4.4c)$$

$$\ln\left(MHY_{it}\right)-\ln\left(MHY_{it-1}\right)=\eta_{mh}\left[a_4\ln\left(OTM_{it}\right)+b_4\ln\left(INM_{it}\right)+c_4\ln\left(EMP_{it}\right)+\sum_{k_4=1}^{K_4}x_{4k_4}\ln\left(\mathrm{X}_{k_4it}^{mh}\right)\right]-\eta_{em}\ln\left(MHY_{it-1}\right)$$

$$\ln\left(MHY_{it}\right)\ =\beta_{41}\ln\left(OTM_{it}\right)+\beta_{42}\ln\left(INM_{it}\right)+\beta_{43}\ln\left(EMP_{it}\right)+\sum_{k_4=1}^{K_4}\gamma_{4k_4}\ln\left(\mathrm{X}_{k_4it}^{mh}\right)+(1-\eta_{mh})\ln\left(MHY_{it-1}\right) \qquad (4.4d)$$

Equations (4.4a) – (4.4d) are the structural equations which constitute the basic simultaneous-equations model in this study. Thus, the general form of the model to be estimated and extended (to accommodate spatial effect) in subsequent sections can be given by:

$$\left\{\begin{array}{l}\ln\left(INM_{it}\right)\ =\beta_{11}\ln\left(OTM_{it}\right)+\beta_{12}\ln\left(EMP_{it}\right)+\beta_{13}\ln\left(MHY_{it}\right)+\sum_{k_1=1}^{K_1}\gamma_{1k_1}l\,\mathrm{n}\left(\mathrm{X}_{k_1it}^{in}\right)+(1-\eta_{in})\ln\left(INM_{it-1}\right)+u_{it}^{in}\\ \ln\left(OTM_{it}\right)\ =\beta_{21}\ln\left(OTM_{it}\right)+\beta_{22}\ln\left(EMP_{it}\right)+\beta_{23}\ln\left(MHY_{it}\right)+\sum_{k_2=1}^{K_2}\gamma_{2k_2}l\,\mathrm{n}\left(\mathrm{X}_{k_2it}^{ot}\right)+(1-\eta_{ot})\ln\left(OTM_{it-1}\right)+u_{it}^{ot}\\ \ln\left(EMP_{it}\right)\ =\beta_{31}\ln\left(OTM_{it}\right)+\beta_{32}\ln\left(INM_{it}\right)+\beta_{33}\ln\left(MHY_{it}\right)+\sum_{k_3=1}^{K_3}\gamma_{3k_3}\ln\left(\mathrm{X}_{k_3it}^{em}\right)+(1-\eta_{em})\ln\left(EMP_{it-1}\right)+u_{it}^{em}\\ \ln\left(MHY_{it}\right)\ =\beta_{41}\ln\left(OTM_{it}\right)+\beta_{42}\ln\left(INM_{it}\right)+\beta_{43}\ln\left(EMP_{it}\right)+\sum_{k_4=1}^{K_4}\gamma_{4k_4}\ln\left(\mathrm{X}_{k_4it}^{mh}\right)+(1-\eta_{mh})\ln\left(MHY_{it-1}\right)+u_{it}^{mh}\end{array}\right\}\ldots(4.5)$$

Spatial Model

Equations in (4.5) are estimated using data collected for cross sectional observations on aggregate spatial units such as counties. Such data sets, however, are likely to exhibit a lack of independence in the form of spatial autocorrelation. When there are no strong *a priori* theoretical reasons to believe that interdependences between spatial units arises either due to the spatial lags of the dependent variables or due to spatially autoregressive error terms, the standard approach is to model the system with both effects included (Anselin, 2003). The spatial autoregressive model with both the spatial lag and spatial error effects can be expressed as:

$$\left\{\begin{aligned}
\ln\left(INM_{it}\right) &= \beta_{11}\ln\left(OTM_{it}\right)+\beta_{12}\ln\left(EMP_{it}\right)+\beta_{13}\ln\left(MHY_{it}\right)+\sum_{k_1=1}^{K_1}\gamma_{1k_1} l\,\mathrm{n}\left(\mathrm{X}_{k_1 it}^{in}\right)+(1-\eta_{in})\ln\left(INM_{it-1}\right)\\
&\quad +\lambda_{11}W\ln\left(INM_{it}\right)+\lambda_{12}W\ln\left(OTM_{it}\right)+\lambda_{13}W\ln\left(EMP_{it}\right)+\lambda_{14}W\ln\left(MHY_{it}\right)+\rho_1 Wu_{it}^{in}+\varepsilon_{\mathrm{it}}^{\mathrm{in}}\\
\ln\left(OTM_{it}\right) &= \beta_{21}\ln\left(OTM_{it}\right)+\beta_{22}\ln\left(EMP_{it}\right)+\beta_{23}\ln\left(MHY_{it}\right)+\sum_{k_2=1}^{K_2}\gamma_{2k_2} l\,\mathrm{n}\left(\mathrm{X}_{k_2 it}^{ot}\right)+(1-\eta_{ot})\ln\left(OTM_{it-1}\right)\\
&\quad +\lambda_{21}W\ln\left(OTM_{it}\right)+\lambda_{22}W\ln\left(OTM_{it}\right)+\lambda_{23}\mathrm{W}\ln\left(EMP_{it}\right)+\lambda_{24}W\ln\left(MHY_{it}\right)+\rho_2 Wu_{it}^{ot}+\varepsilon_{it}^{ot}\\
\ln\left(EMP_{it}\right) &= \beta_{31}\ln\left(OTM_{it}\right)+\beta_{32}\ln\left(INM_{it}\right)+\beta_{33}\ln\left(MHY_{it}\right)+\sum_{k_3=1}^{K_3}\gamma_{3k_3}\ln\left(\mathrm{X}_{k_3 it}^{em}\right)+(1-\eta_{em})\ln\left(EMP_{it-1}\right)\\
&\quad +\lambda_{31}W\ln\left(EMP_{it}\right)+\lambda_{32}W\ln\left(OTM_{it}\right)+\lambda_{32}\mathrm{W}\ln\left(INM_{it}\right)+\lambda_{34}W\ln\left(MHY_{it}\right)+\rho_3 Wu_{it}^{em}+\varepsilon_{\mathrm{it}}^{\mathrm{em}}\\
\ln\left(MHY_{it}\right) &= \beta_{41}\ln\left(OTM_{it}\right)+\beta_{42}\ln\left(INM_{it}\right)+\beta_{43}\ln\left(EMP_{it}\right)+\sum_{k_4=1}^{K_4}\gamma_{4k_4}\ln\left(\mathrm{X}_{k_4 it}^{mh}\right)+(1-\eta_{mh})\ln\left(MHY_{it-1}\right)\\
&\quad +\lambda_{41}\ln\left(MHY_{it}\right)+\lambda_{42}W\ln\left(OTM_{it}\right)+\lambda_{43}\mathrm{W}\ln\left(INM_{it}\right)+\lambda_{44}W\ln\left(EMP_{it}\right)+\rho_4 Wu_{it}^{mh}+\varepsilon_{it}^{mh}
\end{aligned}\right\} \quad ...(4.6)$$

where β, γ, λ, and ρ are unobserved parameters $u_{it}^{in}, u_{it}^{ot}, u_{it}^{em}$, and u_{it}^{mh} are vectors of disturbances, and $\boldsymbol{\varepsilon}_{it}^{in}, \boldsymbol{\varepsilon}_{it}^{ot}, \boldsymbol{\varepsilon}_{it}^{em}$ and $\boldsymbol{\varepsilon}_{it}^{mh}$ are vectors of innovations. K_j, $j = 1,...,4$ represents the number of exogenous variables included in the jth equation. **W** is a row standardized weight matrix

Equations in (4.6) can also be more conveniently written in matrix notation as follows:

$$\mathbf{Y}=\mathbf{YB}+\mathbf{X\Gamma}+\mathbf{WY\Lambda}+\mathbf{U} \tag{4.7a}$$

With

$$\mathbf{U}=\mathbf{WUC}+\mathbf{E}$$

Or more compactly as:

$$\begin{aligned}\mathbf{y} &= \mathbf{B}^{*}\mathbf{y}+\mathbf{\Gamma}^{*}\mathbf{x}+\mathbf{u},\\ \mathbf{u} &= \mathbf{C}^{*}\mathbf{u}+\boldsymbol{\varepsilon}\end{aligned} \tag{4.7b}$$

where $\mathbf{y} = vec\left(\mathbf{Y}\right)$, $\mathbf{x} = \mathrm{vec}\left(\mathbf{X}\right)$, $\mathbf{u} = \mathrm{vec}\left(\mathbf{U}\right)$, and $\boldsymbol{\varepsilon} = \mathrm{vec}\left(\mathbf{E}\right)$,

$$\mathbf{B}^{*} = \left[\left(\mathbf{B}'\otimes\mathbf{I}\right)+\left(\mathbf{\Lambda}'\otimes\mathbf{W}\right)\right] \quad and \quad \mathbf{\Gamma}^{*} = \left(\mathbf{\Gamma}'\otimes\mathbf{I}\right)$$

Assuming that $\mathbf{I}_{nG}-\mathbf{B}^{*}$ and $\mathbf{I}_{nG}-\mathbf{C}^{*}$ are nonsingular matrices with $\left|\rho_j\right|<1$, j = 1,...,4, the system in equation (4.7b) can be expressed in its reduced form as:

$$\begin{aligned}\mathbf{y} &= \left(\mathbf{I}_{nG}-\mathbf{B}^{*}\right)^{-1}\left(\mathbf{\Gamma}^{*}\mathbf{x}+\mathbf{u}\right),\\ \mathbf{u} &= \left(\mathbf{I}_{nG}-\mathbf{C}^{*}\right)^{-1}\boldsymbol{\varepsilon}\end{aligned} \tag{4.7c}$$

Since the innovations are assumed to be independently and identically distributed, that is, $E\left(\boldsymbol{\varepsilon}\right)=\mathbf{0}$ and $E\left(\boldsymbol{\varepsilon\varepsilon}'\right)=\boldsymbol{\Sigma}\otimes\mathbf{I}_n$, the means and variance covariance matrices of the disturbance terms **u**, and the endogenous variables **y**, are given, respectively, as follows:

$$
\begin{aligned}
&E\left(\mathbf{u}\right)=0; \qquad E\left(\mathbf{uu}'\right)=\boldsymbol{\Omega}_{\mathbf{u}}=\left(\mathbf{I}_{nG}-\mathbf{C}^{*}\right)^{-1}\left(\boldsymbol{\Sigma}\otimes\mathbf{I}_n\right)\left(\mathbf{I}_{nG}-\mathbf{C}^{*\prime}\right)^{-1}\\
&E\left(\mathbf{y}\right)=\left(\mathbf{I}_{nG}-\mathbf{B}^{*}\right)^{-1}\left(\boldsymbol{\Gamma}^{*}\mathbf{x}\right); \qquad E\left(\mathbf{yy}'\right)=\boldsymbol{\Omega}_{\mathbf{y}}=\left(\mathbf{I}_{nG}-\mathbf{B}^{*}\right)^{-1}\boldsymbol{\Omega}_{\mathbf{u}}\left(\mathbf{I}_{nG}-\mathbf{B}^{*\prime}\right)^{-1}
\end{aligned}
\tag{4.8}
$$

The endogenous variables as well as the disturbances are, therefore, seen to be correlated both spatially and across equation, and furthermore will generally be hetroskedastic. In this study, the spatial units are counties and each county has only a small number of neighbors and, in turn, it is only a neighbor to a small number of counties. The weights matrix **W** is a row standardized sparse matrix and hence the row and column sums of the weights matrix is bounded in absolute values. It is also assumed that $\left(\mathbf{I}_n-\rho_j\mathbf{W}\right)$, j = 1,...,4 and $\left(\mathbf{I}_n-\mathbf{B}^{*}\right)^{-1}$ are bounded uniformly in absolute values, which imply that $\boldsymbol{\Omega}_{\mathbf{u}}$ and $\boldsymbol{\Omega}_{\mathbf{y}}$ are also bounded uniformly as it can easily be seen from the relations in equation (4.8). Thus, the degree of correlation between the elements of **u** and **y** are limited, which is a necessary condition for all large sample analysis (see Kelejian and Prucha, 1998, 2004).

By imposing exclusion restrictions on the system in equation (4.7a), the spatial autoregressive model can also be reformulated as follows:

$$
\begin{aligned}
\mathbf{y}_j &= \mathbf{Z}_j\boldsymbol{\delta}_j+\mathbf{u}_j,\\
\mathbf{u}_j &= \rho_j\mathbf{W}\mathbf{u}_j+\boldsymbol{\varepsilon}_j, \quad j=1,\ldots,4
\end{aligned}
\tag{4.9}
$$

Where $\mathbf{Z}_j=\left(\mathbf{Y}_j,\mathbf{X}_j,\mathbf{WY}_j\right)$ and $\boldsymbol{\delta}_j=\left(\boldsymbol{\beta}'_j,\boldsymbol{\gamma}'_j,\boldsymbol{\lambda}'_j\right)'$

with $\mathbf{Y}_j,\mathbf{X}_j$ and $\mathbf{WY}_j$ representing the matrices of observations on the endogenous variables, exogenous variables and the spatially lagged endogenous variables that appear in the jth equation respectively.

Data Type and Sources

The data used for the empirical analysis represent the 418 Appalachian counties, which were collected and compiled from County Business Patterns, Bureau of Economic Analysis, Bureau of Labor Statistics, Current Population Survey Reports, County and City Data Book, U.S. Census of Population and Housing, U.S. Small Business Administration, and Department of Employment Security. County-level data for employment, gross in-migration, gross out-migration and median household income were collected for 1990 and 2000. In addition, data for a number of control variables were collected for 1990 from various sources (see table 1 for the data description).

Estimation Issues

The model given in (4.9) is estimated using generalized spatial two stage least squares (GS2SLS) and generalized spatial three stage least squares (GS3SLS) procedures for data from Appalachian counties for 1990-2000. To determine whether a linear or log-linear specification is appropriate for this model, a PE test is undertaken following Kmenta (1986). The test indicates that the log-linear specification is more preferred to the linear form for all equations. Thus, the model is specified in log-linear form with two modifications involving the measurement of the explanatory variables. First, the natural log formulation is dropped for the explanatory variables that can take negative or zero values. Second, lagged 1990 values are used for all of the explanatory variables to avoid simultaneity bias. Hausman's (1978) specification test is also used to test for the endogeniety of the several of the explanatory variables and, accordingly, the four equations are appropriately chosen. Tests for over-identifying restrictions also suggest a proper specification of the model. The presence of spatial autocorrelation in the disturbances is tested using Moran's I test for models with endogenous regressors as suggested in Anselin and Kelejian (1997). All equations of the model show spatial error dependence.

The GS2SLS and GS3SLS procedures are done in a three and a four step routines, respectively. The first three steps are common for both routines. In the first step, the parameter vector consisting of betas, lambdas and gammas $[\beta', \lambda', \gamma']$ are estimated by two stage least squares (2SLS) using an instrument matrix **N** that consists of a subset of $\mathbf{X}, \mathbf{WX}, \mathbf{W^2X}$, where **X** is the matrix that includes all control variables in the model, and **W** is a weight matrix. The disturbances for each equation in the model are computed by using the estimates for $[\beta', \lambda', \gamma']$ from the first step. In the second step, these estimates of the disturbances are used to estimate the autoregressive parameter rho (ρ) for each equation using Kelejian and Prucha's (2004) generalized moments procedure. In the third step, a Cochran-Orcutt-type transformation is done by using the estimates for rhos from the second step to account for the spatial autocorrelation in the disturbances. The GS2SLS estimators for $[\beta', \lambda', \gamma']$ are then obtained by estimating the transformed model using $[\mathbf{X}, \mathbf{WX}, \mathbf{W^2X}]$ as the instrument matrix.

Although the GS2SLS takes the potential spatial correlation into account, it does not utilize the information available across equations because it does not take into account the potential cross equation correlation in the innovation vectors $\boldsymbol{\varepsilon}_{it}^{in}, \boldsymbol{\varepsilon}_{it}^{ot}, \boldsymbol{\varepsilon}_{it}^{em}$ and $\boldsymbol{\varepsilon}_{it}^{mh}$. The full system information is utilized by stacking the Cochran-Orcutt-type transformed equations (from the second step) in order to estimate them jointly. Thus, in the fourth step the GS3SLS estimator of $[\beta', \lambda', \gamma']$ is obtained by estimating this stacked model. The GS3SLS estimator is more efficient relative to GS2SLS estimator. We only report the FGS3SLS estimates.

Results and Discussion

The GS3SLS parameter estimates of the system given in (4.9) are reported in Table 2. The estimated equations explained 94.5, 96.8, 85.7 and 67.5 percent of the variations in employment growth (small business growth), gross in-migration, gross out-migration and median household income growth, respectively. These values are better than results from many studies on cross-sectional analyses of this sort (Deller, *et al*., 2001; Henry *et al*., 1997, Boarnet, 1994). The parameter estimates are mostly consistent with the theoretical expectations. The contemporaneous effects with respect to employment (small business growth), gross in-migration, gross out-migration and median household income are highly significant indicating the existence of very strong feed-back simultaneity among the dependent variables of the empirical model. The results also show strong spatial autoregressive lag and spatial cross-regressive lag simultaneities. Most of the coefficients for the lagged dependent variables are small, although insignificant, indicating short adjustment lags.

EMP Equation:

The results indicate that the level of employment in a county is strongly dependent on the level of contemporaneous gross in-migration, gross out-migration, and median household income. Each of these variables, in turn, is strongly affected by the level of contemporaneous employment. The coefficient for INMG00, for example, is positive and significant at the one percent level. The coefficient for EMP00 in the INMG equation is also positive and significant at the one percent level. These indicate that counties with high levels of in-migration are favorable for small business growth and the growth in small business further encourages in-migration into the counties. But note that the attractive effect of business growth (employment) is more than the effect of gross in-migration on employment as indicated by the level of the coefficients on the respective variables. This is in consistent with the Todaro-thesis of rural-urban migration. A single job opening encourages more than one migrant. Similarly, the interdependence between the level of employment and gross out-migration is very strong but negative. The coefficient for OTMG00 is negative and statistically significant at the one percent level. The coefficient for EMP00 in the OTMG equation is also negative and statistically significant at the one percent level. This means counties with out-migration have factors that discourage small business growth and absence of small business growth, in turn, encourages out-migration. Now again, the contemporaneous effects of EMP on OTMG are stronger than that of OTMG on EMP as indicated by their respective coefficients. The results also show strong positive feed-back simultaneity between EMP and MHY. This is indicated by the positive and statistically significant coefficient for MHY in the EMP equation and the statistically significant coefficient for EMP in the MHY equation. This interdependence is consistent with economic theory and research results in the literature. Note, however, that the attractive effect of median household income on small business growth (employment) is weaker than that of small business growth on median household income.

The results suggest a positive and significant parameter estimate for the spatial autoregressive lag variable (WEMP00). This coefficient represents the spatial autoregressive simultaneity and indicates that the level of employment in a given county tends to spillover to neighboring counties and has a positive effect on their levels of employment. The results also show a positive and significant parameter estimate on the spatial cross-regressive variable with respect to gross out-migration (WOTMG00) indicating that an increase in gross out-migration in neighboring counties tends to encourage business (employment) in a given county. This is possible because the out-migrants from neighboring counties may end up in the county providing the capital and labor that are required for business expansion. Our results also show negative and significant spatial cross-regressive effects with respect to gross in-migration and median household income. This would mean that increases in gross in-migration into and median household income in neighboring counties tend to discourage business (employment) in a given county. This is consistent with economic theory because an increase in income in neighboring counties encourages firms and people to migrate to the neighboring counties in search of markets and jobs. But the migrating firms and people take the capital and labor as well as the skills that are necessary for business expansion out of the given county leading to the decline in employment and business growth in that county. These are important from a policy perspective as they indicate that business growth or the level of employment in one county has positive spillover effects to business growth in neighboring counties. The results are also important from an economic perspective because the significant spatial autoregressive lag and spatial cross-regressive lags effects indicate that EMP00 does not depend only on characteristics within the county, but also on that of its neighbors. Hence, spatial effects should be tested for in empirical work involving employment, gross in- and out-migration as well as household income. The model specification in this study also incorporates spatially autoregressive spatial process (effect) in addition to the spatial lag in the dependent variables. The results in Table 2 suggest a positive parameter estimate for rho1 indicating that random shocks into the system with respect to employment do not only affect the county where the shocks originated and its neighbors, but create positive shock waves across Appalachia.

The coefficient on the initial employment level (EMP) is positive but not statistically significant. This coefficient estimate enables us to calculate the speed of adjustment, eta ($\boldsymbol{\eta}$), in the system with respect to employment. A high speed of adjustment ($\boldsymbol{\eta}$=0.993) is found which indicates a short period of adjustment towards long-run equilibrium. According to this result, about 99 per cent of the equilibrium rate of growth in employment was realized during the ten-year period. This tends to indicate the existence of dynamic stability in the system with respect to employment, but the result is inconclusive since the coefficient estimate on the lag dependent variable is statistically insignificant. Thus, it is difficult to claim that conditional convergence with respect to employment has occurred. Note also that the speed of adjustments ($\boldsymbol{\eta}$) for INMG, OTMG and MHY are 0.362, 0.988, and 0.977, respectively, and the interpretation is similar to that of EMP.

To control for agglomeration effects, the model includes measure of population statistics such as the percentage of population between 25 and 44 years old (POP25_44). The results show that POP25_44 has positive and significant effects on EMP00. This result is consistent with the previous studies (Acs and Armington, 2004) which indicates that a growing population increases the demand for consumer goods and services, as well as the pool of potential entrepreneurs which encourages business formation. This result is important from a policy perspective. It indicates that counties with high population concentration are benefiting from the resulting agglomerative and spillover effects that lead to localization of economic activities, in line with Krugman's (1991a, 1991b) argument on regional spillover effects. However, contrary to the theoretical expectations, the initial human capital endowment as measured by the percentage of adults (over 23 years old) with college degree (POPCD), showed a negative result. The logical interpretation of this result is that the jobs created in Appalachia during the study period might not require high education levels.

Establishment density (ESBd), which is the total number of private sector establishments in the county divided by the total county's population, is included in our model to capture the degree of competition among firms and crowding of businesses relative to the population. The coefficient for ESBD is negative and significant indicating that Appalachian region has reached the threshold where competition among firms for consumer demands crowds businesses. According to the results, ESBd is associated with low level of employment (business growth), indicating that firms seem to have exhausted localization and agglomeration economies of scale.

The coefficient for the variable representing the median value of housing (MVH) is positive and significant at the one percent level. This result indicates that high value of owner occupied housing is positively associated with business formation in Appalachia. This result is consistent with the theoretical expectation that high value of housing is an indication that there is a capacity to finance new business by potential entrepreneurs, either by using the house as collateral for loan application or it is an indication of availability of personal financial resources to start new business.

The coefficients for variables MANU and WHRT are both positive and significant at the one percent levels. These results indicate that counties with initial higher percentage of their labor force employed in manufacturing and in wholesale and retail trades showed higher growth rate in business formation than other counties.

In-Migration - INMG Equation:

The results from the INMG equation also indicate that the level of gross in-migration into a county is strongly dependent on the level of contemporaneous level of employment, gross out-migration, and median household income. These interdependences are explained by the highly statistically significant coefficients for the endogenous variables of the model. The feed back simultaneity between gross out-migration and gross in-migration is positive and very strong, indicating that counties that are characterized by high gross out-migration were also counties of high migration destinations during the study period. This is possible because out-migrants and in-migrants could be people with different labor market characteristics. Besides, a growing share of in-migrants in a county reflects a growing share of migration-prone residents, which is likely to increase out-migration from the county. The migration literature also indicates that migrants in one period are more likely than non-movers to move in subsequent periods.

The interdependence between gross in-migration and median household income is negative and strong. This indicates that high income counties are associated with low in-migration. This could be due to the fact that some migrants prefer low income locations. Clark and Hunter (1992), for example, found that movers in their early 20s as well as migrants 35 years and older prefer low-income locations. Besides, Knapp and Graves (1989) suggest that higher income locations may be associated with low levels of amenities that discourage people from migrating in.

Turning to the spatial autoregressive lag and spatial cross-regressive lag effects, the coefficient for the spatial autoregressive lag variable fails to be significant indicating the absence of spatial autocorrelation with respect to in-migration. The coefficients on the spatial cross-regressive lag variables with respect to employment (WEMP00) and out-migration (WOTMG00), however, are negative and statistically significant at the one per cent level. This indicates that in-migration in one county is negatively associated with employment and out-migration in neighboring counties. Neighboring counties' household incomes, on the other hand, have positive effects on the level of in-migration in a given county as indicated by the positive and statistically significant parameter estimate for WMHY00.

The results in Table 2 suggest a positive parameter estimate for rho2 indicating that random shocks into the system with respect to gross in-migration do not only affect the county where the shocks originated and the effects on the neighbors, but create positive shock waves across Appalachia.

The initial period population size (POPs) showed a positive and strong effect on in-migration into a given county. The positive and statistically significant coefficient for POPs is an indication that people migrate to areas (counties) with high concentrations of population. Note also that the coefficient for POPs in the out-migration equation is negative and statistically significant at the one per cent level, indicating that factors in counties with high population concentrations discourage out-migration and vice versa. These two results suggest that Appalachian counties that are characterized by small and dispersed communities have been losing people during the study period. Since Appalachia is dominated by small communities, this finding is important from a policy perspective. The out-migration of people, mostly the young and more educated, from the small communities would mean the erosion of the income and property tax bases that provide the major sources of revenue to finance local public services such as schools, hospitals, and other infrastructures. These effects, in turn, increase the tax price per remaining person for any level of public spending and consequently, the per capita cost of providing local public services for the community at large increases. Over time, the quality and quantity of local public services in the community deteriorates, and further out-migration results. The results of this study also show that an increase in the tax price per capita discourages in-migration and encourages out-migration into and out of a given county. This is indicated by the significant positive and negative coefficients for the EXTAX variable, which is derived by dividing the per capita tax revenue by local government expenditure per capita, in the in-migration and out-migration equations, respectively. However, note that a declining population not only increases the per capita cost of providing local public services but also constrains the expansion and growth of small business by limiting the supply of labor and the demand for small business products. Low quality and quantity of public services also reduces the earning capacity of residents and discourages small business growth and employment. The ultimate result is the perpetuation of poverty and underdevelopment in Appalachia. The policy implication of these findings is that counties and communities need to create enabling economic environments for firms and people to stay to do business and work in their jurisdictions. The significant spatial interdependence would also mean that neighboring counties may need to pool their resources in their efforts to promote growth.

Out-Migration - OTMG Equation:

The results from the out-migration equation also show similar trends. The feed-back simultaneities are very strong. Gross out-migration is associated negatively with contemporaneous level of employment and positively with contemporaneous in-migration and median household income, as indicated by statistically significant (all at the one percent level) coefficients for EMP00, INMG00, and MHY00. The results also show a strong positive spatial autoregressive lag effect as indicated by the statistically significant coefficient for WOTMG00 and positive spatial cross-regressive lag effect with respect to employment as indicated by the statistically significant coefficient for WEMP00. This result suggests that out-migration in one county is associated with high levels of out-migration and employment in neighboring counties. The positive spatial autoregressive lag effect shows that spatial clustering with respect to out-migration exists in Appalachia. Counties with declining populations are losing their population together with their neighbors. The policy implication of this finding is that counties with declining population may need to pool their resources to deal with their problems.

The results in Table 2 also suggest a positive parameter estimate for rho3 indicating that random shocks into the system with respect to gross out-migration do not only affect the county where the shocks originated, but has effects on the neighbors and create a positive shock waves across Appalachia.

The impact of home ownership on out-migration is positive and significant which is not consistent with theoretical expectations. Normally, owing a house would be expect to decrease the propensity to migrate due to the transaction costs and its liquidity of real estate in locations of economic distress. Investing in housing of your own may also reflect a decision to stay in the area of current residence for long. Contrary to these theoretical expectations, the empirical results, however, show a positive and statistically significant (at the one per cent level) coefficient for OWHU. This result indicates that home ownership is positively associated with out-migration in Appalachia during the study period. This result, however, reflects the reality in Appalachia. Home ownership in Appalachia is positively associated with the level of economic distress during the study period - home ownership is higher in distressed counties (76 percent) and lower in attainment counties (72 percent); higher in central Appalachia than in Northern or Southern sub regions (more developed); and Appalachia non-metro areas had higher ownership rates (76 percent) than its metro areas (72 percent) (Pollard, 2003). Thus, this result further reflects that the direction of migration in Appalachia during the study period was from small dispersed and distressed communities.

Median Household Income - MHY Equation:

Similar to what we have in the other equations, the estimates from the MHY equation show the existence of significant feed-back simultaneity, spatial autoregressive lag simultaneity and cross-regressive lag simultaneities. The contemporaneous effect with respect to the level of employment on median household income is positive and statistically significant at the one percent level. This result indicates that high level of median household income is positively associated with high level of employment which is consistent with the expectations of economic theory. The contemporaneous effect with respect to the level of out-migration on the level of median household income is also positive and statistically significant at the one percent level. This result suggests that median household income increases with out-migration. This, in turn, would mean that the average income of the out-migrants is lower than the median income of the non-movers. The contemporaneous effect with respect to in-migration on the level of median household income is negative and statistically significant at the one percent level. This result also indicates that median household income in a given county is negatively associated with the level of in-migration to that county. This, in turn, suggests that the average income of the in-migrants is lower that the median income of the non-movers. These two results, thus, suggest, compared to the non-movers, that the movers are poor. Based on this result, it is, therefore, possible for one to claim that the population movement in Appalachia during the study period is on average for economic reasons.

The spatial autoregressive lag effect is positive and statistically significant at the one percent level, indicating that the level of median household income in a given county is positively affected by the level of median household income in neighboring counties. This strong spatial spillover effect is an indication that there is clustering of counties in Appalachia on the bases of their median household incomes. The exploratory spatial data analysis on the same data set (not shown here) indication that most of the low income counties are clustered in Central

Appalachia, whereas the high income counties are clustered, mostly, around larger cites, in the Northern and Southern Appalachian sub regions. There is also a significant negative spatial cross-regressive lag effect with respect to the level of employment, indicating that the level of median household income in a given county is negatively associated with the level of employment in neighboring counties. This could be as a result of the fact that an increase in the employment level in neighboring counties may indicate favorable conditions for business in those counties that attract both people and firms even from the given county, which ultimately leads to the loss of income in that county. The spatial cross-regressive lag effect with respect to the level of out-migration is also negative and statistically significant at the one percent level. Since the direction of migration, as indicated above, is out from distressed counties, this result could indicate that neighbors of distressed counties are also distressed, a further indication of the clustering of poverty in Appalachia.

The results in Table 2 also suggest a negative parameter estimate for rho4 indicating that random shocks into the system with respect to median household income do not only affect the county where the shocks originated and the effects on the neighbors, but create negative shock waves across Appalachia.

Turning to the conditioning variables in the MHY equation, the empirical results indicate that the level of median household income is positively and significantly affected by the initial population size (POPs), the percentage of families with female family householder (FHHF), and the social capital index (SCIX). POPs is positively associated with MHY due to the beneficial effects of agglomeration economies of firm location. A growing population captures the extent to which counties are relatively attractive to migrants and a growing population increases the demand for consumer services which in turn leads to growth in business and employment, which are themselves sources of income to the county. The coefficient for the index of social capital (SCIX) is also positive and significant indicating that counties with high level of social capital increase the wellbeing of their communities. The result is consistent with the expectation of economic theory. The positive effects of the FHHF on MHY are not consistent with the theoretical expectations. Perhaps, those households could be located and have greater opportunities for employment in the higher income counties whereas in the poorer counties there would be few opportunities for employment even at low wage. However, the proportion of female family householder per se is not what is important. The characteristics and the earning capacity, which have more to do with many personal, social and economic factors, is what matters. Thus, in a priori we cannot claim that FHHF should be inversely related to MHY.

Note also that the proportion of the population 25 years and above with four years college degree (POPCD) was not significant in the EMP and MHY equations. Human capital theory postulates that entrepreneurship is related to educational attainment and work experience. People with more educational skills tend to form businesses and also have more probability of getting and securing higher paying jobs. Long periods of lack of economic opportunity in Appalachia, however, have led to the continued out-migration of the more educated and skilled portion of the population. Thus, the insignificant effects of POPCD in both the employment and the median household income equations could be an indication of the result of this long term trend.

Conclusions and Policy Implications

As previously indicated, the main objective of this study is to test the hypotheses that (1) business growth (the level of private employment used as its proxy), gross in-migration, gross out-migration and median household income are interdependent and are jointly determined by regional covariates; (2) business growth, gross in-migration, gross out-migration and median household income in one county are conditional upon the respective variables in neighboring counties, and (3) the existence of dynamic stability or convergence towards long-run equilibrium in the system defines the interdependences among these variables. To test these hypotheses, a spatial simultaneous equations growth equilibrium model is developed. Feasible Generalized

Spatial Three-Stage Least Squares (FGS3SLS) estimates of the coefficients of the model parameters are obtained by estimating the model using county-level data covering the 418 Appalachian counties for the 1990-2000. Empirical evidence indicates support for all the three hypotheses. In particular, very strong feed-back simultaneities among the dependent variables of the model exist. The results also indicate significant spatial autoregressive lag simultaneities as well as spatial cross-regressive lag simultaneities in all of the equations of the model. These results indicate that the dependent variable of a given equation in the model is not only conditioned by what is happening in a given county, but also by what is happening to the dependent variables in neighboring counties as well. The study results also indicate the presence of spatial correlation in the error terms. This implies that a random shock into the system spreads across the region. A policy implication of these spatial interdependence and significant spatial spillover effects is that neighboring counties may need to pool and integrate their resources to encourage the positive spatial spillover effects and jointly mitigate the negative spatial spillover effects. The speed of adjustment parameters was also calculated and the results show the existence of short lag adjustments in the system. Although, these results suggest that dynamic stability exists in the system with respect to the three dependent variables (EMP00, OTMG00, and MHY00), it is difficult to claim that it really exists in the system. This is because the coefficients for the lagged dependent variables upon which the speed of adjustments are calculated are insignificant (lower than ten percent level).

The results also show that the direction of migration in Appalachia during the study period was out from low populated, dispersed and distressed counties. Small communities are losing both people and firms. The implication is that such communities will not be able to sustain average quality of life for their residents in the long-run.

References

Acs, Z. J. and C. Armington, 2004, "The Impact of Geographic Differences in Human Capital on Service Firm Formation Rates," *Journal of Urban Economics*, forthcoming.

Anselin, L. 2003, "Spatial Externalities, Spatial Multipliers and Spatial Econometrics," *International Regional Science Review*, **26**(2):153-166.

_______, and Harry H. Kelejian, 1997, " Testing for Spatial Error Autocorrelation in the Presence of Endogenous Regressors," *International Regional Science Review*, **20**:153-182.

Aronsson, T., J. Lundberg and M. Wikstrom**,** 2001, "Regional Income Growth and Net Migration in Sweden 1970-1995," *Regional Studies,* forthcoming.

Barkley, D.L., M.S. Henry, and S. Bao, (1998), "The Role of Local School Quality and Rural Employment and Population Growth," *Review of regional Studies,* **28** (summer): 81-102.

Boarnet, M.G., (1994), "An Empirical Model of Intra-metropolitan Population and Employment Growth," *Papers in Regional Science*, **73** (April):135-153.

Carlino, O.G. and E.S. Mills, 1987, "The Determinants of County Growth," *Journal of Regional Science* **27**(1):39-54.

Clark, D. and C.A. Murphy, 1996, "Countywide Employment and Population Growth: An Analysis of the 1980s," *Journal of Regional Science* **36**(2):235-256.

_______, and W.J. Hunter, 1992, "The Impact of Economic opportunity, Amenities and Fiscal Factors on Age Specific Migration Rates," *Journal of Regional Science*, **32**(3):349-365.

Cushing, Brain, and Cynthia Rogers, 1996, "Income and Poverty in Appalachia," *Socio-Economic Review of Appalachia (Papers Commissioned by the Appalachia Regional Commission)*, Regional Research Institute, West Virginia University.

Deavers, Kenneth L., and Robert A. Hoppe, 1992, "Overview of the Rural Poor in the 1980s," in Cynthia Duncan (ed.), *Rural Poverty in America*, Westport, CT: Greenwood Publishing Group Inc.

Deller, Steven C., Tsung-Hsiu Tsai, David W. Marcouiller, and Donald B.K. English, 2001, "The Role of Amenities and Quality of Life in Rural Economic Growth," *Journal of American Journal Agricultural Economics,* **83**(3):352-365.

Dilger, Robert Jay, and Tom Stuart Witt, 1994, "West Virginia's Economic Future, " in Robert Jay Dilger and Tom Stuart Witt (eds.), *West Virginia in the 1990s: Opportunities for Economic Progress*, Morgantown: West Virginia University Press.

Duffy, N.E., 1994, "The Determinants of State Manufacturing Growth Rates: A Two-Digit-Level Analysis," *Journal of Regional Science,* **34**(May):137-162.

Duffy-Deno, K.T., 1998, "The Effect of Federal Wilderness on County Growth in the Inter-mountain Western United States," *Journal of Regional Sciences*, **38**(February): 109-136.

Duffy-Deno, Kevin T. and Randall W. Eberts, 1991, "Public Infrastructure and Regional Economic Development: A Simultaneous Equations Approach," *Journal of Urban Economics* **39**:329-343.

Edmiston, Kelly D., 2004, "The Net Effect of Large Plant Locations and Expansions on Country Employment," *Journal of Regional Science*, **44**(2):289-319.

Ekstrom, B., and F. L. Leistritz, 1988, *Rural community decline and revitalization: An annotated Bibliography*. New York: Garland Publishing. Inc.,

Greenwood, M. J. and G.L. Hunt, 1984 "Migration and Interregional Employment Redistribution in the United States," *American Economic Review* **74**(5):957-969.

Greenwood, Michael J., Gary L. Hunt and John M. McDowel, 1986, "Migration and Employment Change: Empirical Evidence on Spatial and Temporal Dimensions of the Linkage," Journal of Regional Science, 26(2):23-234.

Hamalainen, Kari, and Petri Bockerman, (2004), "Regional Labor Market Dynamics, Housing, and Migration," *Journal of Regional Science*, **44**(3):543-568.

Harrison, B. and S. Kanter, 1978, "The Political Economy of States: Job-Creation and Business Incentives," *American Institute of Planning Journal*, (October):425-435.

Hausmann, J., 1978, "Specification Tests in Econometrics," *Econometrica*, **46**:1251-1271.

Hayness, Ada F., 1997, *Poverty in Central Appalachia: Underdevelopment and Exploitation*, New York: Garland Publishing, Inc.

Henry, M.S., D.L. Barkley, and S. Bao, 1997, "The Hinterland's Stake in Metropolitan Growth: Evidence from Selected Southern Regions," *Journal of Regional Science*, **37**(August): 479-501.

Henry, M.S., B. Schmitt, K Kristensen, D.L. Barkley, and S. Bao, 1999, "Extending Carlino-Mills Models to Examine Urban Size and Growth Impacts on Proximate Rural Areas," *Growth and Change*, **30**(Summer):526-548.

Kelejian, Harry H. and Ingram R. Prucha, 1998, "A Generalized Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Spatial Autoregressive Disturbances," *Journal of Real Estate, Finance and Economics*, **17**(1):99-121.

_____________, 2004, “Estimation of Simultaneous Systems of Spatially Interrelated Cross Sectional Equations,” *Journal of Econometrics*, **118**: 27-50.

Kmenta, J., 1986, *Elements of Econometrics*, New York: Macmillan.

Knapp, T. A., and P. E. Graves, 1989, “On the Role of Amenities in Models of Migration and Regional Development,” *Journal of Regional Science*, **29**(1):71-87.

Krugman, P. 1991a, “Increasing Returns and Economic Geography,” *Journal of Political Economy*, **99**:483-499.

_____, 1991b, *Geography and Trade*, Cambridge, MA: MIT Press.

Lewis, David J., Gray L. Hunt and Andrew J. Plantigna, 2002, “Does Public Land Policy Affect Local Wage Growth,” *Growth and Change* **34**(1):64-86.

MacDonald, John F., 1992, “Assessing the Development Status of Metropolitan Areas,” in Edwin S.Mills and John F. MacDonald (ed.), *Sources of Metropolitan Growth*, New Brunswick, NJ: Center for Urban Policy Research.

Maggar, Sally W., 1990, “Schooling, Work Experience and Gender: The Social Reproduction of Poverty in Mining Regions,” Research Paper 9014, Regional Research Institute, Morgantown: West Virginia University Press.

Mills, E.S., and R. Price, 1984, “Metropolitan Suburbanization and central City Problems,” *Journal Urban Economics*, **15**:1-17.

Pulver, G. C., 1989, “Developing a Community Perspective on Rural Economic Development Policy.” *Journal of Community Development Society* 20:1-4.

Rupasingha, A. and S.J. Goetz, 2003, “ County Amenities and Net Migration,” Paper Presented in WRSA Annual Meeting, Rio Rico, Arizona, Feb 26-March 1.

Steinnes, Donald N. and Walter D. Fisher, 1974, “An Econometric Model of Intra-urban Location,” *Journal of Regional Science* **14**:65-80.

White, Nancy E. and Thomas A. Knapp, 1994, “A Dynamic Model of Migration with Public Sector Attributes,” *Papers in Regional Science* 73(3):331-340.

Table 1: Descriptive statistics

Variable Code	Variable Description	Mean	Std Dev	Minimum	Maximum
EMP00	Employment 2000	0.112	0.427	-0.927	7.212
INMG00	In-migration 2000	7.133	0.983	4.277	10.619
OTMG00	out-migration 2000	9.934	0.304	5.778	10.681
MHY00	Median Household income 1989	7.236	0.331	6.492	10.897
WEMP00	Spatial Lag of EMP00	0.103	0.258	-0.459	1.614
WINMG00	Spatial lag of INMG00	7.131	0.547	5.645	8.703
WOTMG00	Spatial Lag of OTMG00	9.939	0.163	9.052	10.382
WMHY00	Spatial lag of MHY00	7.237	0.170	6.821	7.890
AREA	County Area in square miles	6.009	0.748	1.099	7.277
POPs	Population 1990	10.297	0.948	7.877	14.106
POP25_44	Percent of population between 25 -14 years old 1990	3.380	0.077	2.785	3.745
FHHF	percent of female householder, family householder, 1990	2.322	0.203	1.811	3.188
POPCD	Persons 25 years and over, % high school or higher, 1990	4.100	0.171	3.570	4.468
OWHU	Owner-Occupied Housing Unit in percent, 1990	4.325	0.076	3.867	4.473
MVH	Median Value of owner occupied housing 1990	10.737	0.263	9.668	11.676
MCRH	Median Contract of Rent of Specified Rent-Occ.,1990	5.641	0.206	4.942	6.358
UNEMP	Unemployment rate 1990	2.154	0.348	1.224	3.246
MANU	% employed in manufacturing 1990	3.184	1.040	0.788	21.000
WHRT	% employed in wholesale and retail trade 1990	2.948	0.650	2.163	15.600
PCPTAX	Property tax per capita 1990	5.524	0.616	3.912	7.363
SCIX	Social Capital Index 1987	-0.593	0.960	-2.527	5.645
NAIX	Natural Amenities Index 1990	0.143	1.159	-3.720	3.550
HWD	Highway Density 1990	0.690	0.404	-0.339	2.632
ESBD	Establishment density 1990	2.928	0.335	1.874	4.093
EXTAX	General Expenditure per unit tax paid	0.843	0.514	-0.984	2.608
EMP90	Employment 1990	7.188	0.990	4.575	10.907
INMG90	In-migration 1990	7.037	0.977	4.500	10.550
OTMG90	out-migration 1990	0.107	0.324	-1.095	4.990
NMHY90	Median Household income 1989	10.419	0.240	9.697	11.127

Note: *All the variables are expressed in log terms except SCIX, and NAIX*

Table2: Feasible Generalized Spatial Three-Stage Least Squares(FGS3SLS) Estimation Results

	EMP Equation		INMG Equation		OTMG Equation		MHY Equation	
VARIABLE	*Coefficient*	*t-statistic*	*Coefficient*	*t-statistic*	*Coefficient*	*t-statistic*	*Coefficient*	*t-statistic*
CONSTANT	-1.279*	-1.871	-6.037***	-6.515	4.856***	6.800	-3.666***	-2.616
EMP00			0.737***	11.776	-0.682***	-19.128	0.785***	9.286
INMG00	0.167***	5.113			0.337***	14.558	-0.251***	-4.485
OTMG00	-0.415***	-8.048	0.938***	9.977			0.579***	4.676
MHY00	0.101**	2.306	-0.318***	-6.586	0.144***	4.295		
WEMP00	0.760***	26.155	-0.182***	-2.648	0.454***	12.044	-0.660***	-7.400
WINMG00	-0.066***	-4.356	0.025	1.018	-0.024	-1.378	0.011	0.346
WOTMG00	0.476***	8.541	-0.260***	-2.815	0.342***	5.471	-0.393***	-3.187
WMHY00	-0.248***	-5.617	0.134*	1.868	-0.066	-1.316	0.872***	12.496
AREA			0.018	1.289	-0.033***	-3.379		
POPs			0.302***	8.556	-0.242***	-10.492	0.200***	4.129
POP25_44	0.241***	3.526						
FHHF							0.306***	4.567
POPCD	-0.006	-0.328					0.098	0.557
OWHU					0.380***	4.287		
MVH	0.108***	2.926					0.091	1.306
MCRH			-0.005	-0.071				
UNEMP			-0.050*	-1.869	-0.009	-0.496	-0.014	-0.496
MANU	0.080***	10.971						
WHRT	0.253***	12.590						
PCPTAX	0.003	0.204						
SCIX							0.049**	2.009
NAIX	-0.003	-0.735	0.001	0.194	-0.006	-1.415		
HWD	-0.018	-1.291						
ESBd	-0.054***	-3.335						
EXTAX			-0.132***	-4.999	0.178***	12.786		
EMP90	0.007	0.224						
INMG90			0.638***	16.951				
OTMG90					0.012	0.314		
MHY90							0.023	0.568
RHO (ρ)	0.081		0.034		0.125		-0.499	
SIGM (σ)	0.008		0.022		0.01		0.054	
R^2	0.945		0.968		0.857		0.675	
	418		418		418		418	
ETA (η)	0.993		0.362		0.988		0.977	

Note: **, **, and *** denote statistical significance level at 10 percent, 5 percent and 1 percent respectively.*