ISSN 1084-1695

Aging Studies Program Paper No. 15

When Random Group Effects are Cross-Correlated: An Application to Elderly Migration Flow Models

Karen Smith Conway and Andrew J. Houtenville*

Maxwell Center for Demography and Economics of Aging

Center for Policy Research
Maxwell School of Citizenship and Public Affairs
Syracuse University
Syracuse, New York 13244-1020

October 1998

Support for the Aging Studies Program Series is provided by grant number P20-AG12837 from the National Institute on Aging.

Abstract

Incorporating random group effects has proven important to making correct statistical inferences about factors that only vary across groups. We note that it is possible to have more than one random effect in models using cross-sectional data and that these random effects could be correlated, unlike in the typical panel data situation. Extending the standard multiple random effects model in this way is greatly simplified by using the two-step estimator we develop. Our application to an elderly migration flow model provides an intuitive example of cross-correlated random group effects and demonstrates the ease of our estimator, as well as highlighting the empirical importance of controlling for random effects.

1. Introduction

Researchers tend to view random effects models, especially two-way random effects models, as applying mainly to situations using panel data. In such cases, the two random effects refer to time and cross-section effects, and one can sensibly assume that these effects are uncorrelated. However, as pointed out by Moulton (1986), Case (1991) and others, random effects can also prove important in data drawn from a population with a grouped structure. For instance, cross-sectional microeconomic data frequently contain repeated observations from the same geographic location such that a "location" effect could exist. It is also possible to have more than one grouping structure present in the data and therefore more than one random effect. And, unlike the panel data case, these random group effects *could be* correlated.

One example in health economics is patient treatment or outcome data that could contain both a doctor and a hospital effect. To the extent that doctors choose the hospitals at which they practice, these two effects could be correlated (e.g., high quality doctors choose high quality hospitals, or doctors with preferences for certain kinds of treatment choose hospitals that excel in that treatment). Similarly, in regulation studies, the choice or outcome of a regulatory action may be a function of both the individual enforcing the regulation and the regulatory climate in which the action takes place. Thus, both an "enforcer" and a location effect are present and these two effects may be correlated (enforcers are more zealous in locations where they receive greater support). In our empirical application, we propose yet another situation of two cross-correlated random effects in a migration flow model where there is both a destination and an origin effect. Here, the random influences that cause individuals to leave a certain state (origin effect) should be correlated with those that draw individuals to that state (destination effect). We are the first to introduce random effects into a migration flow model and to explore whether they are correlated.[1]

Another possibility is that the random effects are correlated *within* the group, which may occur because there are further subgroupings within a grouping. In our examples above, the random effects for doctors in the same practice or the random effects for enforcers who work for the same agency could be correlated. In these instances, the correlation arises because the groups, doctors and enforcers, can be further grouped by practice or agency. In our application to a migration flow model, states in the same region could have random origin or destination effects that are correlated. This is a general problem that may also afflict models with only one random effect, which are far more common.

Of course, one could always capture these influences by estimating them as fixed effects (i.e., including dummy variables for each grouping or subgrouping structure). However, this treatment typically precludes the estimation of the effects of observed factors that only vary across groups; e.g., one cannot estimate group varying variables both a hospital fixed effect and the effect of certain hospital characteristics.[2] When these group-invariant factors are of interest, one must specify the unobserved group effects as random, and yet Moulton (1986) shows that these are the precise instances in which properly controlling for random effects is so important. In particular, using microeconomic data he reveals that ignoring group effects (i.e., using OLS) biases downward the estimated standard errors of the group-invariant coefficients, thereby artificially inflating their statistical significance. We find a similar result in our empirical application.

Incorporating group effects is therefore essential to making correct statistical inferences about coefficients on factors that only vary across groups and, unless one has panel data, these effects must be specified as random. As illustrated by the examples above, these random effects could be correlated, both *among* levels within a group and *across* groupings (when more than one random effect exists). However, the standard approach, grounded in panel data applications, is to assume they are uncorrelated.[3] Ignoring cross-correlation not only results in possibly incorrect statistical inferences, it also neglects a potentially interesting piece of information, as our examples reveal. In the next section, we construct a general multi-random effect model that relaxes these assumptions and devise a relatively simple two-step estimator with which to estimate it. We then follow with an application to a migration flow model that provides an intuitive example of cross-correlated random group effects and demonstrates the ease of our estimator, as well as the empirical importance of properly controlling for random effects. Because it deals with one of the fastest growing and potentially most mobile groups in society—the elderly—our empirical results should be of interest as well.

2. A General Model of Multi-Group, Cross-Correlated Random Effects

The examples discussed above and our empirical application are all special cases of a general model, which can be written using Moulton's (1986) notation as

$$y = X\beta + Z_1\delta_1 + \ldots + Z_c\delta_c + u, \tag{1}$$

where X is an n x p matrix of known constants, β is a p vector of unknown parameters, Z_i is an n x q_i matrix of dummy variables for each of the q_i levels of the ith grouping, δ_i is a q_i x 1 vector of random variables with assumed mean zero and covariance $\sigma_i^2 I$, c is the number of *grouping variables* or effects, and u is the typical white noise error. For instance, in the health economics example, n is the number of patients (or observations), the number of grouping variables (c) is two—hospital and doctor—and q_1 and q_2 are the number of hospitals and doctors, respectively, in the data. Moulton (1986) assumes that the random effects are uncorrelated, both across groupings (i.e., $E[\delta_i \, \delta_j'] = 0, \forall \, i \neq j$) and among levels within a group (i.e., $E[\delta_{il} \, \delta_{im}] = 0, \forall \, i = 1,..,$ c groupings and $l \neq m$ levels). The covariance matrix of the composite error is then simply

$$V = \sigma_u^2 I + \sum_{i=1}^{c} \sigma_i^2 Z_i Z_i'. \qquad \textbf{(2)}$$

This covariance matrix is therefore relatively straightforward and, as Moulton (1986) notes, the model can be estimated via maximum likelihood. Note that the "unbalanced" nature of the "panel," that is, that the number of levels, q_i, is not the same across all c groupings (nor is the number of observations belonging to each level), makes the usual simplifying transformation using "between" and "within" estimates more difficult to implement even when c equals 2.[4]

We now extend this framework in two ways. First, we explicitly acknowledge the presence of group-invariant variables by respecifying δ_i from equation (1) to have both an observed and an unobserved component,

$$\delta_i = W_i \alpha_i + \varepsilon_i, \quad i = 1, \ldots, c, \qquad \textbf{(3)}$$

where W_i is q_i x k_i matrix of group-invariant variables, α_i is a k_i vector of unknown parameters and ε_i is the purely random component of δ_i. Returning to the health economics example, W_1 is a matrix of k_1 hospital characteristics and α_1 is the vector of coefficients of those characteristics. Substituting equation (3) into (1) yields

$$y = X\beta + Z_1 W_1 \alpha_1 + \ldots + Z_c W_c \alpha_c + Z_1 \varepsilon_1 + \ldots + Z_c \varepsilon_c + u. \qquad \textbf{(4)}$$

Writing the model using equations (1), (3) and (4) places it in the same general framework as that of Amemiya (1978) in his development of a simplifying two-step estimator. We apply the insights from Amemiya (1978) in devising our 2-step estimator, and return to this issue shortly.

The second extension is allowing the random elements of δ to be correlated *across groupings* and *among levels within a grouping*. To deal with correlation across groupings, which is our primary focus, we assume that $E[\varepsilon_i \varepsilon_j']$ is a q_i x q_j matrix Δ_{ij} containing at least some nonzero elements. Note that Δ_{ij} need not be a square matrix because the number of levels, q, need not be the same across grouping structures (e.g., the number of hospitals need not equal the number of doctors). To be completely general we also allow for correlation among levels so that Δ_{ii} need not equal $\sigma_i^2 I$, as assumed in equation (2), and instead must only have variances on the diagonals and be positive definite. The covariance matrix of the composite error is now

$$V \;=\; \sigma_u^2 I \;+\; \sum_{i=1}^{c} \sum_{j=1}^{c} Z_i \Delta_{ij} Z_j'. \qquad \textbf{(5)}$$

Correctly estimating the model requires specifying and getting estimates of V, which is clearly now a much more complicated task. Specifying the Δ_{ij}'s may be relatively straightforward depending on the situation. In both the health economics and regulatory examples, the two random effects are likely only correlated across groupings if they refer to the same general location (e.g., doctors and hospitals in the same area). Similarly, in our migration flow model the destination and origin effects are only assumed to be correlated when they refer to the same state.[5] As for correlation among levels, doctors in the same practice, enforcers in the same agency or states in the same region may have random effects that are correlated.[6] Note that in these instances there is actually a subgrouping of levels that causes the correlation; e.g., doctors may be further grouped by their practice.

One method of estimation is to first estimate the random effects model in the typical way (assuming no correlation) and then use the estimated random components, $\hat{\varepsilon}$, to estimate the Δ_{ij}'s.[7] Equation (4) can then be estimated using feasible GLS. Or, one could use maximum likelihood estimation. Both of these procedures require the inversion of the potentially large (n x n) and complicated covariance matrix written in (5). The two-step estimator we devise next greatly simplifies estimation.

2.1 A Two-Step Estimator

In developing our two-step estimator we borrow heavily from the insights and general framework of Amemiya (1978). For a general model, Amemiya (1978) demonstrates the algebraic equivalence between two different estimators—one obtained by efficiently estimating the entire equation and the other by efficiently estimating a second equation in which the estimated regression coefficients from the first are the dependent variables. He then devises a feasible estimator for a special case of the model which involves panel data and random slope coefficients that vary over cross-sectional unit.[8] We first explain his general result and show how our model fits within his general model. We then extend the insights from his feasible estimator in deriving a two-step estimator for our model.

The general model Amemiya considers is

$$y = X\beta + Z\delta + u \quad \text{and} \tag{6}$$

$$\delta = W\alpha + \varepsilon, \tag{7}$$

where X, Z and W are matrices of known constants, β and α are vectors of unknown parameters and u and ε are vectors of unobservable random variables that are uncorrelated with each other, have zero means and general covariance matrices Λ and Ω, respectively. Substituting (7) into (6) yields

$$y = X\beta + ZW\alpha + Z\varepsilon + u. \tag{8}$$

He derives the BLU estimator for α from this equation, which must take account of the messy error structure. He next considers a two-step estimator for α and shows that it is algebraically equivalent. The first step efficiently estimates δ from equation (6). (This could be the OLS estimator if *u* is the typical white noise error with covariance matrix $\sigma^2 I$). The second step uses the estimated values of δ as dependent variables in efficiently estimating α via equation (7). Note that this estimator must take account of the error, ε, as well as the additional error, $v \equiv \hat{\delta} - \delta$, that comes from using estimated values for δ. Amemiya (1978) shows that these two estimators for α are algebraically equivalent, yet the two-stage estimator is much simpler computationally.

Our model fits easily within this general framework. Our matrix *Z* is an *n* x *G* matrix equal to $[Z_1 \ Z_2 \ ... \ Z_c]$ where *G* is the sum of all q_i levels over all *c* groupings. The Z_i's are the matrices of dummy variables defined in equation (1), such that it is the intercept that is allowed to be random rather than some other coefficient. δ is a *G* x 1 vector of the estimated grouping-level effects.[9] *W* is a *G* x *H* matrix equal to diag(W_1, W_2, ..., W_c) where *H* is the number of explanatory variables, k_i, summed over all *c* groupings. α is the *H* x 1 vector of parameters to be estimated, and ε is the *G* x 1 vector of the purely random effects. The covariance matrix of ε consists of Δ_{ii} on the q_i block diagonals and Δ_{ij} on the q_i x q_j off-diagonals, or

$$Cov(\varepsilon) = \Omega = \begin{bmatrix} \Delta_{11} & \Delta_{12}' \\ \Delta_{12} & \Delta_{22} \end{bmatrix} \quad \textbf{(9)}$$

for the case of two random effects. The covariance of the composite disturbance in the second stage is therefore *Cov*(ε) + *Cov*(v), the latter of which is estimated from the first stage (and is the variances and covariances of the estimated dummy variable coefficients).

However, because we have two or more sets of dummy variables (*c* > 1), the two-step estimator Amemiya (1978) proposes and Borjas and Sueyoshi (1994) extend is not estimable. In essence, this two-step estimator requires estimating fixed effects for every level (*q*) of every grouping structure (*c*), which is not possible. The matrix *Z* is singular; e.g., one cannot estimate a fixed effect for each origin and each destination in our migration flow model. In the case of random *slope* coefficients (Amemiya) or *one* random effect (Borjas and Sueyoshi), this problem is resolved by suppressing the general coefficient.[10] Because we have *c* random effects, *c* restrictions are required in order to estimate these fixed effects. We impose the restriction that the q_i effects for each grouping structure sum to zero by omitting the dummy variable for the last level in each grouping, and we retain the general intercept. Thus, for each grouping structure only q_i - 1 dummy variables are included.

Given these restrictions *Z* is no longer singular and given our assumptions about *u*, consistent estimates of $\hat{\beta}$ and $\hat{\delta}$ of equation (6) can be obtained using OLS. In other words, the coefficients on group-*variant* variables (β) are consistently estimated using a fixed effects estimator in the first stage. The estimated group effects, $\hat{\delta}$, are then regressed on the relevant characteristics via equation (7) using FGLS, taking account of the covariance structure of ε discussed above, plus the error introduced by using estimated values of δ instead of the true parameters (i.e., the fact that $Cov(\hat{\delta})$ is not "nice"). However, our restrictions have altered the interpretation of these group effects. The estimated dummy coefficient for grouping *i*, level *j*, $\hat{\delta}_{ij}$, is really the difference between the effects of level j and the omitted level, q_i. Specifically,

$$\hat{\delta}_{ij} = \delta_{ij} - \delta_{iq_i} + v_{ij} , \qquad \textbf{(10)}$$

for *i*=1...*c* groupings and *j*=1...q_i-1 levels within each grouping, and *v* is the error resulting from using the estimated values of δ, which have nonconstant variances and nonzero covariances. Substituting in from equation (7),

$$\hat{\delta}_{ij} = (W_{ij} - W_{iq_i})\alpha_i + \varepsilon_{ij} - \varepsilon_{iq_i} + v_{ij} . \qquad \textbf{(11)}$$

This can be written more generally as

$$\hat{\delta} = \tilde{W}\alpha + \tilde{\varepsilon} + v , \qquad \textbf{(12)}$$

where $\hat{\delta}$ is a vector containing the *G-c* estimated grouping-level effects (or dummy coefficients), and all other terms are as defined for equation (7) except that the omitted levels' observed variables and random effect are being subtracted from *W* and ε respectively (and hence the use of the tilde). Also the omitted level no longer appears in $\hat{\delta}$, *W*, ε or *v*, such that the number of rows is now *G-c*. We write out all of these matrices (both for equations (6)-(8) and (12)) for the special case of our migration flow model in the technical appendix.

Clearly, the covariance of $\tilde{\varepsilon}$ is more complicated than that of ε, written in equation (9). The diagonal blocks now contain an *additional* σ_i^2 because all of the random effects for group *i* contain that of the omitted level and therefore have a covariance. Of course, depending on the form of correlation among levels, there may be other modifications too. Likewise, the off-diagonal blocks depend on the original form of the cross-correlation, the Δ_{ij}'s. As shown in our application, these transformations may not be that complex. If there is correlation among levels, carefully choosing which level to omit may reduce the complexity (e.g., choose a level that is *not* correlated with any others). Even though this new covariance matrix is more complicated, it is still a much smaller matrix to invert (*G-c* x *G-c*) than that written in equation (5) (*n* x *n*). For instance, in our empirical application (using the 48 contiguous states) *G-c* is 94, and *n* is 2,256!

In sum, the two-step estimator involves first estimating the dependent variable as a function of all of the group-*variant* variables *X* and *G-c* grouping-level dummy variables via OLS to obtain $\hat{\beta}$ and $\hat{\delta}$. The estimated dummy variable coefficients, $\hat{\delta}$, and their covariances are then used in a second step in which the dummy coefficients are regressed on the group-invariant characteristics of that group-level minus those of the omitted level. Both elements of the composite error term, $\tilde{\varepsilon} + v$, are messy. $\tilde{\varepsilon}$ is complicated by the correlation between ε_i and ε_j and because by eliminating one dummy variable $\tilde{\varepsilon}$ is now the difference between each level and the omitted level. v comes from using the estimated dummy variable coefficients from the first stage, which have nonconstant variances and nonzero covariances, as do any set of estimated regression coefficients. This composite error necessitates estimating the second stage equation (equation 11 or 12) by FGLS in order to get asymptotically efficient estimates of and correct standard errors for α.

2.2 Estimating the Covariances and Testing for Cross-Correlation

The elements of the Δ_{ij}'s (and $\tilde{\Delta}_{ij}$'s) can be estimated with the second stage OLS residuals. However, the residuals have several components, at least in finite samples. Specifically, the cross-product of the second stage residuals is equal to

$$e_{il}e_{jm}=(\tilde{\varepsilon}_{il} + v_{il} + \tilde{W}_{il}(\alpha_i-\hat{\alpha}_i))(\tilde{\varepsilon}_{jm} + v_{jm} + \tilde{W}_{jm}(\alpha_j-\hat{\alpha}_j)), \quad \textbf{(13)}$$

where the first subscript refers to the grouping and the second refers to the level within the grouping structure. The first component of both residuals is the random group effect. The second component is the sampling error from the first stage estimation—the fact that $\hat{\delta}$ and not δ is the dependent variable in the second stage. Likewise, the third component is the sampling error of the second stage estimation.

However, the second and third components of the residual are zero asymptotically, because the parameters δ and α are consistently estimated using OLS in the first and second stages. As with any FGLS estimator, our random effects estimator only contains asymptotic properties which means that these two components can be safely ignored. Thus, a valid estimate of the l, m element of $\tilde{\Delta}_{ij}$ is the cross-product of the corresponding residuals, or

$$\lim_{n\to\infty} E[e_{il}e_{jm}] = E[\tilde{\varepsilon}_{il}\tilde{\varepsilon}_{jm}] = \tilde{\Delta}_{ij_{lm}}. \quad \textbf{(14)}$$

This is essentially the feasible estimator that Amemiya proposes (1978, equation (23), pp. 795); however, because he assumes that the elements of the covariance matrix are constant across the cross-sectional units, his estimator uses the mean over the cross-sections. When more structure is placed on the Δ_{ij}'s, as in our application, we also suggest using the mean of all squares or cross-products that have the same asymptotic expectation.

In principle, one can adjust for the finite components of the residual, although the gains are unclear. For instance, Borjas and Sueyoshi (1994, p. 179) modify Amemiya's estimator by taking account of the second component of the residual, *v*. Ignoring the last component in equation (13) (as Borjas and Sueyoshi do), the expected value of the residual cross-products is

$$E[e_{il}e_{jm}] = E[\tilde{\varepsilon}_{il}\tilde{\varepsilon}_{jm}] + E[v_{il}v_{jm}] + E[\tilde{\varepsilon}_{il}v_{jm}] + E[\tilde{\varepsilon}_{jm}v_{il}], \qquad \textbf{(15)}$$

where the last two terms are approximately zero because ε and *v* are assumed independent. Fortunately we have estimates of the second term from our first stage estimates—$E[v_{il}v_{jm}]$ is the covariance between the corresponding estimated dummy variable coefficients. This suggests a slightly different estimator,

$$\tilde{\Delta}_{ij_{lm}} = e_{il}e_{jm} - Cov(\hat{\delta}_{il}\hat{\delta}_{jm}). \qquad \textbf{(16)}$$

This is a more general form of the estimator proposed by Borjas and Sueyoshi (1994). In their model, there is only one random effect and it is uncorrelated among levels such that $Cov(\varepsilon)$ is simply $\sigma_{\varepsilon}^2 I$. Thus, they recommend using the sum of the squared second stage OLS residuals minus the sum of the estimated dummy variable coefficient variances from the first stage, all divided by the appropriate degrees of freedom.

In addition, the third component of the residual, the second stage sampling error, could be taken into consideration. However, this is much more difficult because this component is correlated with the other two and the resulting covariances will be functions of all of the elements of $Cov(\tilde{\varepsilon})$ and $Cov(v)$, as well as the matrix $\tilde{W}$. For this reason, we recommend either the asymptotic estimator suggested by equation (14) or making the finite adjustment for *v* as in equation (16). We derive both estimators for our empirical application in the next section.

Also of interest is whether the random group effects are indeed correlated, for if they are not then the model collapses into a standard random effects model.[11] Again, the specific test for cross-correlation depends on the exact form specified and whether one wants an individual or a joint test of cross-correlation. A Lagrange multiplier test is appropriate to test an individual correlation parameter,

$$\lambda_{LM} = m\rho_{ij}^2 , \qquad \textbf{(17)}$$

where *m* is the number of estimated second stage residual cross-products used to calculate σ_{ij} (and the Pearson correlation coefficient ρ_{ij}). The statistic is distributed chi-squared with one degree of freedom. This test is a modified version of the Lagrange multiplier test for cross-sectional correlation discussed in Greene (1997, p. 661), the latter of which can also be used for a more general test of Δ_{ij} equals 0.

3. Empirical Application—An Elderly Migration Flow Model

The usual specification of migration flow models is based on the "gravity model," which posits that states with large populations pull in and push out more people and that the distance between two states dampens these effects. This specification is typically written as

$$\ln M_{ij} = \beta_0 + \beta_1 \text{distance}_{ij} + \alpha^o \ln \text{Pop}_i + \alpha^d \ln \text{Pop}_j + \mu_{ij} , \qquad \textbf{(18)}$$

for $i = (1,...,m)$ origin states and $j = (1,..., i\text{-}1, i\text{+}1,...m)$ destination states, and where M_{ij} is the number of people who move from state *i* to state *j*, and μ_{ij} is a assumed to be white noise error.[12] This model fits nicely into the discussion above. There are two groupings, origins and destinations, denoted by superscripts *o* and *d*, respectively. Distance is a group-variant variable, and the natural log of population is a group-invariant variable because it remains the same for given group (for instance, all flows out of state *i*). α^o and α^d are the effects of the natural logs of the origin's and destination's populations, respectively.

Going beyond the "gravity model," an economic model of migration behavior has the utility maximizing individual weighing the costs and benefits associated with his or her current location and other potential locations (see Greenwood (1975) and Graves and Linneman (1979)). Graves and Knapp (1988), Dresher (1994) and Conway and Houtenville (forthcoming) incorporate the public sector into the elderly's maximization calculus. The elderly are prime candidates for Tiebout behavior ("voting" and "shopping" with one's feet) as retirement could precipitate such a move both by changing their preferences and income and by freeing them from labor market concerns. The elderly's preferences for certain kinds of publicly provided goods (e.g., Medicaid and other health expenditures as opposed to education expenditures) differ in a systematic way from the preferences of other groups of individuals. Their sources of income (pensions and interest income), as well as their expenditure patterns (dis-saving), imply that certain taxes are less burdensome than others.

Examining the flow of elderly migrants from one location to another is one way to explore elderly migration.[13] Five studies explore these flows, of which only two incorporate aspects of the public sector, Voss, Gunderson and Manchin (1988) and Conway and Houtenville (1998).[14] The other three migration flow studies include a general measure of taxes only as a proxy for or as part of a cost-of-living index (McLeod, Parker, Serow and Rives 1984, Serow, Charity, Fournier and Rasmussen 1986, and Fournier, Rasmussen and Serow 1988). All explicitly investigate whether the effects of destination and origin characteristics are symmetric and find that the destination and origin coefficients are often of the *same* sign, sometimes statistically significantly so (see Voss, Gunderson and Manchin 1988 for a lengthy discussion of this problem).

To incorporate this broader notion of migration behavior and to allow independent variables and possibly correlated random effects, the “gravity model” is expanded such that

$$\ln M_{ij} = X_{ij}\beta + W_i\alpha^o + W_j\alpha^d + \mu_{ij}, \qquad \textbf{(19)}$$

where

$$\mu_{ij} = \varepsilon_i^o + \varepsilon_j^d + u_{ij}. \qquad \textbf{(20)}$$

Referring back to equation (4) reveals this as a special case of our general model, written here for one observation, *ij*. Again, there are two groupings in the data, origin and destination ($c = 2$), and 2,256 (48 times 47) possible migration flows ($n = 2{,}256$). In addition to the common intercept and distance, X_{ij} contains any other factors that vary within groups such as whether state *i* and *j* share a border. W_i is a *k* vector of group-invariant characteristics of state *i*, such as the natural log of population and public sector variables. W_i is assumed to contain the same factors for both the origin and destination states and so α^o and α^d are now the corresponding *k* vectors of coefficients for the origin and destination state characteristics, respectively. ε_i is the unobserved characteristics and any other random influences of state *i* and u_{ij} is the typical white noise error. The technical appendix discusses in greater detail how our model fits into the framework of equations (6)-(8) and writes out many of the matrices. Notice also that the number of levels within each grouping is equal—$q_o = q_d = 48$. However, it is not a “balanced panel”; whereas each origin has 47 possible destinations, they are not the *same* 47 destinations because of the necessary omission of the origin state.[15]

In this model, the two random effects, ε^o and ε^d may be correlated in several ways. They may be correlated *across groupings* for the same state—the random influences that cause people to choose state *i* for their destination are likely (negatively) correlated with the influences that cause people to leave state *i*. It is also possible that they may be correlated *among levels*—the random influences that draw people to California (ε_i^d) may be correlated with those that draw people to Oregon or Nevada (ε_j^d). This leads to the possibility of correlation *across groupings and among levels*—the unobserved reasons that people leave California (ε_i^o) could be correlated with the reasons people move to Nevada (ε_j^d). Yet another complication raised by the possibility of among level correlation in our application is the role of spatial correlation. As discussed in Anselin (1988), spatial correlation between locations is a function of how "close" they are and is multi-directional (e.g., California is correlated strongly with Oregon and Nevada, and less strongly to Idaho and Utah, who are *all* also correlated with each other to varying degrees). Properly allowing for *among level* correlation therefore greatly complicates our analysis and hinders our primary purpose of demonstrating how to use our two-step estimator. Thus, we limit ourselves to the consideration of *across grouping* correlation and leave the incorporation of among level correlation in a migration flow model to future research.

Formally, the assumptions we make regarding the two random effects can be written as

$$\varepsilon^o \sim N\left(0,\ \sigma_o^2 I\right),\quad \varepsilon^d \sim N\left(0,\ \sigma_d^2 I\right) \quad \text{and} \quad E\left[\varepsilon_i^o \varepsilon_j^d\right] = 0$$
$$\text{when } i \neq j \quad \text{and} \quad E\left[\varepsilon_i^o \varepsilon_j^d\right] = \sigma_{od} \text{ when } i = j\ . \tag{21}$$

This specification leads to several different correlations across observations. Errors for migration flows with the same destination are correlated, as are those with the same origin. This is the usual case of a two-way random effects model. However, the correlation between ε^d and ε^o means that the errors for migration flows *to* state i are correlated with the errors for flows *from* state *i*. Defining the composite error for $\ln M_{ij}$ as μ_{ij} ($\equiv \varepsilon_i^o + \varepsilon_i^d + u_{ij}$) the possible covariances are

$$E\left(\mu_{ij}\,\mu_{ik}\right) = \sigma_o^2 \qquad \text{(same origin)},$$
$$E\left(\mu_{ij}\,\mu_{kj}\right) = \sigma_d^2 \qquad \text{(same destination)},$$
$$E\left(\mu_{ij}\,\mu_{ki}\right) = \sigma_{od} \qquad \text{(destination the same as } origin\text{)}, \tag{22}$$
$$E\left(\mu_{ij}\,\mu_{ji}\right) = 2\sigma_{od} \qquad \text{(opposite flows)},$$
$$E\left(\mu_{ij}\,\mu_{kl}\right) = 0 \qquad \text{(no state in common)},$$

for $i \neq j \neq k \neq$ l.

3.1 Implementing the Two-Step Estimator

Using the framework set forth in section 2 greatly simplifies matters. Assuming that the origin and destination effects are listed in the same order, the covariance matrix between ε^o and ε^d, Δ_{od}, equals $\sigma_{od}I$. Using the general covariance matrix of the composite error written in equation (5), the covariance matrix for our elderly migration model is simply

$$V = \sigma_u^2 I + \sigma_o^2 Z_o Z_o^{\prime} + \sigma_d^2 Z_d Z_d^{\prime} + \sigma_{od} Z_o Z_d^{\prime} + \sigma_{od} Z_d Z_o^{\prime}. \tag{23}$$

where Z_o and Z_d are matrices of 48 origin and 48 destination dummy variables, respectively.

Although the covariance matrix written in equation (23) looks relatively simple, it is still a nondiagonal *n* x *n* matrix, where *n* = 2,256 state flows in our example. The two-step estimator developed in section 2.1 greatly simplifies things. Corresponding to equation (6), the first stage consists of estimating the migration flows as a function of an intercept, distance, the border dummy, and the origin and destination dummy variables,

$$\ln M_{ij} = X_{ij}\beta + Z_{oi}\delta^o + Z_{dj}\delta^d + u_{ij}. \tag{24}$$

Of course, in order to estimate this model, the last state dummy variable (Wyoming, in our sample) has to be omitted from both Z_o and Z_d. Corresponding to equation (11), the second stage involves estimating the 47 origin and destination dummy coefficients on the state-invariant variables, *W*, taking account of the fact that Wyoming has been omitted,

$$\hat{\delta}_i^o = W_i\alpha^o - W_m\alpha^o + \varepsilon_i^o - \varepsilon_m^o + v_i , \quad and \tag{25}$$

$$\hat{\delta}_i^d = W_i\alpha^d - W_m\alpha^d + \varepsilon_i^d - \varepsilon_m^d + v_{m-1+i} . \tag{26}$$

This was written more generally in equation (12) and the corresponding matrices for our application are written out fully in the technical appendix. Again, the covariance of $\tilde{\varepsilon}$ is more complicated than that of ε, but is fairly straightforward,

$$Cov(\tilde{\varepsilon}) = \begin{bmatrix} \sigma_o^2(I+J) & \sigma_{od}(I+J) \\ \sigma_{od}(I+J) & \sigma_d^2(I+J) \end{bmatrix} , \tag{27}$$

where J is an m-1 square matrix of ones reflecting the fact that all observations share the random effect from the omitted state.

The estimates of σ_o^2, σ_d^2, and σ_{od} are obtained using the second stage (OLS) residuals from the origin and destination dummies, via equation (14), and if one wishes to make the finite correction for v, the estimated (co)variances for the dummy variable coefficients from the first stage as in equation (16). With the finite adjustment these estimators are

$$\hat{\sigma}_o^2 = \frac{1}{2}\frac{1}{m-1}\left(\sum_{i=1}^{m-1} e_i^2 - \sum_{i=1}^{m-1} Var\left(\hat{\delta}_i^o\right)\right) , \tag{28}$$

$$\hat{\sigma}_d^2 = \frac{1}{2}\frac{1}{m-1}\left(\sum_{i=m}^{2m-2} e_i^2 - \sum_{i=1}^{m-1} Var\left(\hat{\delta}_i^d\right)\right) , \tag{29}$$

$$\hat{\sigma}_{od} = \frac{1}{2}\frac{1}{m-1}\left(\sum_{i=1}^{m-1} e_i \, e_{i+m-1} - \sum_{i=1}^{m-1} Cov\left(\hat{\delta}_i^o \, \hat{\delta}_i^d\right)\right) . \tag{30}$$

Note that the finite adjustments, subtracting the estimated first stage variances as in Borjas and Sueyoshi (1994), could lead to negative estimated variances. If this occurs, we recommend omitting this adjustment (as in Amemiya and equation (14)) as it is asymptotically zero anyway.

With our estimate of σ_{od} (and therefore ρ_{od}), we can test for cross-correlation with the Lagrange multiplier test written in equation (17). We can also estimate the second stage equation with FGLS, thereby obtaining estimates for α that control both for the presence of group effects and the possibility they may be cross-correlated. Note again that our two-step estimator requires the inversion of a 94 x 94 matrix, as opposed to the typical way which would require inverting a 2,256 x 2,256 covariance matrix.

3.2 Description of the Data

Table 1 provides a brief description and the mean and standard deviation of each variable used in our analysis. The dependent variable is the natural log of the flow (or number) of elderly individuals (aged 65 and older) who moved out of state *i* and into state *j* between 1985 and 1990. This variable is obtained from the *County-to-County Migration Flow Files* of the 1990 U.S. Census. If there is no migration from state *i* to state *j*, then the natural log of the flow is set to zero.[16]

Recall there are two explanatory variables that are flow specific (X_{ij} in equation (24)) , the distance between state *i* and *j* and a border dummy variable to control for the possibility that information is probably better and the costs of moving lower if the states share a border.[17] Following the "gravity model," the natural log of the state population is included as an origin and destination specific explanatory variable (Z_{io} and Z_{jd} in equation (24)).

The rest of the group-variant explanatory variables fall into four broad categories, cost of living, amenities, government expenditures and tax shares. Our cost-of-living variable is the index created by McMahon (1991) and we expect a high cost-of-living to cause more out-migration ($\alpha^o > 0$) and less in-migration ($\alpha^d < 0$), *ceteris paribus*. Our amenity variables include three measures of climate, as well as median household income and the total number of criminal offenses known to police per 100,000 residents. We disaggregate total state and local government expenditures into several types, all per capita—education expenditures, public welfare expenditures, expenditures on health and hospitals, a measure of Medicaid generosity, and all other expenditures. Our tax variables are state and local property taxes, sales taxes, personal income taxes and all other taxes and sources of revenue (except federal aid and interest), all expressed as a percentage of total state and local expenditures. Because personal income tax structures vary widely over states with respect to the treatment of pension income and social security, its effect on migration may vary across states. We attempt to control for this by including an interaction term that is the product of the personal income tax share and the amount of pension income that is exempt.

Since our dependent variable refers to migration between 1985 and 1990 the exact timing of the impact of the independent variables effect is uncertain. Since all migrants, even those who migrated in 1985, have access to 1984 information our explanatory variables are for the year 1984. In addition, the use of 1984 variables ensures that the elderly migrants do not directly influence the state's population and public sector variables.[18]

Note that with the exception of distance and border, all variables are group-invariant. Thus, group effects are potentially important. And yet, unless one uses panel data, the inclusion of state dummy variables (or fixed state effects) prohibits estimation of the effects of observed state characteristics. Given that Census data is generally only available every ten years, pooling different time periods may be inappropriate, as Mueser (1989) notes. Appropriately specifying random effects is the only viable choice.

3.3 Empirical Results

We estimate our model both with and without the public sector variables since much of the elderly migration literature ignores these variables. These specifications are estimated three ways. We begin by estimating each one with ordinary least squares and with the typical two-way random effects model (using our two-step estimator) in order to explore the impact of controlling for group effects as in Moulton (1986). We reestimate the specification using our two-stage cross-correlated random effects estimator described above if the random effects are found to be cross-correlated.

Table 2 reports the results. The estimated coefficients reveal two salient patterns. First, the point raised by Moulton (1986) is confirmed; ignoring random effects seriously overstates the statistical significance of many of the state variable coefficients. The other salient feature is that the coefficients on the origin variables are almost always the same sign as those for the destination variables, which runs counter to what one would expect based on a simple theoretical model of migration. Recall that this result has been found by many others (e.g., Voss, Gunderson and Manchin (1988)). Including the public sector variables tends to exaggerate this pattern, both by causing some variables to exhibit it (the cost-of-living and heating degree variables) and because these new variables also exhibit it.

Looking at individual coefficients, the distance coefficient is always negative and very statistically significant, and the border, and population variables are always positive and statistically significant. All are consistent with our expectations. The rest of the coefficients, however, are not so nicely behaved, as our expectations of opposite signs are repeatedly violated. For instance, cost of living in the destination state has a negative effect ($\alpha^d < 0$), as expected, but so does the cost of living in the origin state, although it is frequently statistically insignificant (i.e., $\alpha^o \leq 0$). In other words, cost-of-living has a negative effect on in-migration, as expected, but it sometimes also has a negative effect on out-migration, contrary to what is predicted from the theoretical framework. The most significant amenity variable is the crime rate, which has a positive effect on both in- and out-migration. A high crime rate should lead to greater out-migration, but not to greater in-migration. Perhaps this is due to the noisiness of the variable. Fournier, Rasmussen and Serow (1988) use a similar crime variable and find the same odd results. Most of the public sector variables are statistically significant but also have coefficients of the same sign. For instance, education expenditures lead to less in-migration, as expected, but also to less out-migration. The tax share coefficients (except the interaction term) are all positive and statistically significant for both the origin and destination states. This suggests that high tax shares lead to greater out-migration, as expected, but also to greater in-migration.

As revealed at the bottom of Table 2, there is strong evidence of cross-correlated random effects. The estimated variances, σ_o^2 and σ_d^2, are of a significant magnitude and increase as fewer variables are included. The variance of the destination random effects, σ_d^2, is bigger than that of the origin in both cases, which is consistent with the idea that information may be noisier about destination states than origin states. Goldfeld-Quandt tests reject the hypothesis that the two variances are equal in the model with the public sector variables, and fails to reject their equality in the one without them. In addition, Lagrange multiplier tests for the existence of these random effects (Judge et al. 1985, p. 536) all strongly reject the hypothesis of no random effects either jointly or individually.

More importantly, we test for whether the random effects are correlated using the Lagrange multiplier test written in equation (17) and for whether they are perfectly negatively correlated, as suggested by the symmetry argument. Both specifications soundly reject both null hypotheses. In fact, the correlation coefficient between the two is *positive* and actually exceeds 1.0 for the public sector specification if the finite adjustments for the estimated dummy coefficients' covariances are used. For this reason, we are forced to abandon this finite adjustment. The estimates of the variance components with the adjustments are reported in the footnotes to the tables. Without the finite adjustment, the correlation coefficients for the two specifications are 0.743 and to 0.797. Given these high correlations, it is surprising that our results do not change more when we control for cross-correlation. Interestingly, the errors become more highly correlated the *more* explanatory variables that are included. Thus, our results suggest that both the observed and unobserved influences of state characteristics are asymmetric, and indeed tend to move in the *same* rather than opposite direction.

In sum, then, we find that ignoring random effects in migration flow models can substantially overstate the importance of state characteristics. This emphasizes the need to control for state effects, either with fixed effects or random effects, in future research that uses migration flows. We also find the same pattern of asymmetry in the coefficients reported by Voss, Gunderson and Manchin (1988) and others. In particular, many of the public sector and amenity variables appear important to migration flows, but they affect both in-migration and out-migration in the same way. Furthermore, we soundly reject that the random group effects are uncorrelated; instead, we find that they too are positively correlated.

4. Closing Remarks

Incorporating random *group* effects has proven important to making correct statistical inferences about factors that only vary across groups. We note that if there is more than one grouping structure the random effects could be correlated across groupings, unlike in the typical panel data situation. Furthermore, even when there is only one random effect, further sub-groupings in the data may lead the effects to be correlated among levels. Such forms of cross-correlation result in a complicated covariance matrix which could be difficult to invert. We offer a computationally simpler two-step estimator in the spirit of Amemiya (1978) and Borjas and Sueyoshi (1994) and a methodology for testing the cross-correlation.

We illustrate our technique with an intuitive application to elderly migration. By using migration flows, we can compare the characteristics of the destination state with those of the origin to study the factors influencing elderly migration. Our econometric model includes two random group effects, one for the origin and one for the destination, and presumes the two are correlated for the same state. That is, the unobserved influences that cause people to leave a state are likely correlated with those that cause people to move into it. In general, we find strong evidence of random state effects and that ignoring them seriously inflates the empirical importance of many state characteristics. This confirms the findings of Moulton (1986). The random effects also appear to be *positively* correlated and significantly so, which leads to the use of our two-step cross-correlated random effects estimator.

The results of our application suggest that random effects are a critical attribute of migration flow models and that public sector variables are an important factor in elderly migration decisions. Both results warrant future research. More generally, this research reveals that the assumption of no correlation across random effects may be questionable when applied to grouped rather than panel data. Extending the standard multiple random effects model in this way is greatly simplified by the two-step estimator we provide.

Technical Appendix

Our elderly migration flow model fits into general framework of Amemiya (1978), written in equations (6) and (7), when *Z*, *W*, δ, α and ε are defined as below:

$$
Z = \left[\begin{array}{ccccc ccccc}
1 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \\
 & & & & & & & & & \\
0 & 1 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 & 0 & 1 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 1 \\
 & & & & & & & & & \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
 & & & & & & & & & \\
0 & 0 & 0 & \cdots & 1 & 1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 1 & 0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0 & 0 & 0 & \cdots & 0
\end{array}\right]
\begin{array}{l}
 \\ \\ m-1 \text{ rows} \\ \\ \\ \\ \\ \\ \text{m}-1 \text{ rows} \\ \\ \\ \\ \\ \\ \\ \\ \text{m}-1 \text{ rows} \\ \\
\end{array}
,
$$

$$\qquad\qquad \text{m columns} \qquad\qquad \text{m columns}$$

$$
W = \begin{bmatrix}
W_{11} & W_{12} & W_{13} & \cdots & W_{1k} & 0 & 0 & 0 & \cdots & 0 \\
W_{21} & W_{22} & W_{23} & \cdots & W_{2k} & 0 & 0 & 0 & \cdots & 0 \\
W_{31} & W_{32} & W_{33} & \cdots & W_{3k} & 0 & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
W_{m1} & W_{m2} & W_{m3} & \cdots & W_{mk} & 0 & 0 & 0 & \cdots & 0 \\
 & & & & & & & & & \\
0 & 0 & 0 & \cdots & 0 & W_{11} & W_{12} & W_{13} & \cdots & W_{1k} \\
0 & 0 & 0 & \cdots & 0 & W_{21} & W_{22} & W_{23} & \cdots & W_{2k} \\
0 & 0 & 0 & \cdots & 0 & W_{31} & W_{32} & W_{33} & \cdots & W_{3k} \\
\vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\
0 & 0 & 0 & \cdots & 0 & W_{m1} & W_{m2} & W_{m3} & \cdots & W_{mk}
\end{bmatrix} ,
$$

$$
\delta = \begin{bmatrix} \delta_1^o \\ \delta_2^o \\ \vdots \\ \delta_m^o \\ \\ \delta_1^d \\ \delta_2^d \\ \vdots \\ \delta_m^d \end{bmatrix} , \qquad \alpha = \begin{bmatrix} \alpha_1^o \\ \alpha_2^o \\ \vdots \\ \alpha_k^o \\ \\ \alpha_1^d \\ \alpha_2^d \\ \vdots \\ \alpha_k^d \end{bmatrix} \qquad \textit{and} \qquad \varepsilon = \begin{bmatrix} \varepsilon_1^o \\ \varepsilon_2^o \\ \vdots \\ \varepsilon_m^o \\ \\ \varepsilon_1^d \\ \varepsilon_2^d \\ \vdots \\ \varepsilon_m^d \end{bmatrix} .
$$

Substituting these into equations (6) - (8) yields, respectively,

$$
y_{ij} = X_{ij}\beta + \delta_i^o + \delta_j^d + u_{ij} \quad , \tag{1A}
$$

$$
\delta_i^\ell = \alpha_1^\ell W_{i1} + \alpha_2^\ell W_{i2} + \ldots + \alpha_k^\ell W_{ik} + \varepsilon_i^\ell \quad , \tag{2A}
$$

$$
\begin{aligned}
y_{ij} = {} & X_{ij}\beta + \alpha_1^o W_{i1} + \ldots + \alpha_k^o W_{ik} + \alpha_1^d W_{j1} + \ldots \\
& + \alpha_k^d W_{jk} + \varepsilon_i^o + \varepsilon_j^d + u_{ij} \quad ,
\end{aligned} \tag{3A}
$$

where i, j = 1 ... m states and $\ell = o, d$.

It is now clear that Z is a singular matrix and thus equation (1A) cannot be estimated without imposing some restrictions. Deleting the m^{th} and $2m^{th}$ columns of Z imposes the necessary restrictions and results in m-1 origin and m-1 destination dummy variable coefficients, $\tilde{\delta}$, being estimated. However, the interpretation of these $\tilde{\delta}$'s is altered as mentioned in the text. The second stage is now altered to be a function of $\tilde{W}$ and $\tilde{\varepsilon}$:

$$\tilde{W} = \begin{bmatrix} W_{11}-W_{m1} & W_{12}-W_{m2} & \cdots & W_{1k}-W_{mk} & 0 & 0 & \cdots & 0 \\ W_{21}-W_{m1} & W_{22}-W_{m2} & \cdots & W_{2k}-W_{mk} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ W_{m-1,1}-W_{m1} & W_{m-1,2}-W_{m2} & \cdots & W_{m-1,k}-W_{mk} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & W_{11}-W_{m1} & W_{12}-W_{m2} & \cdots & W_{1k}-W_{mk} \\ 0 & 0 & \cdots & 0 & W_{21}-W_{m1} & W_{22}-W_{m2} & \cdots & W_{2k}-W_{mk} \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & W_{m-1,1}-W_{m1} & W_{m-1,2}-W_{m2} & \cdots & W_{m-1,k}-W_{mk} \end{bmatrix}$$

and

$$\tilde{\varepsilon} = \begin{bmatrix} \varepsilon_1^o-\varepsilon_m^o \\ \varepsilon_2^o-\varepsilon_m^o \\ \vdots \\ \varepsilon_{m-1}^o-\varepsilon_m^o \\ \varepsilon_1^d-\varepsilon_m^d \\ \varepsilon_2^d-\varepsilon_m^d \\ \vdots \\ \varepsilon_{m-1}^d-\varepsilon_m^d \end{bmatrix}.$$

Endnotes

[*]The authors are respectively, Associate Professor of Economics at the University of New Hampshire, and Research Associate, Center for Policy Research, The Maxwell School, Syracuse University.

1. Bolduc, Laferriére, and Santarossa (1992) permit origin and destination random effects in a passenger travel flow model applied to synthetic data; however, they maintain the assumption that the two errors are uncorrelated.

2. Unless, of course, one has panel data, in which case these factors vary over time as well as across groups.

3. Moulton (1986) also makes this assumption, as do all other studies of random effect models to the best of our knowledge, except for Bolduc, Laferriére and Santarossa (1992) who allow their random effects to be correlated *among levels* using a spatial autoregressive (SAR) process.

4. Judge et al. (1985, pp. 533-37) discusses the simplifying transformation when c equals 2 and the panel *is* balanced, and Greene (1997) discusses the difficulty of using these transformations with an unbalanced panel when there is only one grouping or one random effect.

5. The role of location here is distinct from that typically assumed in spatial econometrics. Spatial econometrics considers spatial dependence, in which cross sectional units are correlated with one another depending upon their proximity. Defining "close" proximity and modeling a decay effect are issues that arise. In the present study, we treat observations as either belonging to the same general location or not and do not consider the correlation across geographically "close" observations. For further discussion of spatial econometrics, see Anselin (1988) or Case (1991).

6. This is perhaps a fruitful place to build into the model the concept of spatial correlation, as in Bolduc, Laferriére and Santarossa (1992). We discuss this further in the application section.

7. Judge et al. (1985) shows the calculation of the random components for the typical two-way random effects model in a balanced panel.

8. Amemiya allows the coefficients to be correlated for each cross-section, but not across cross-sections. Moulton (1986, p. 393) considers a similar random coefficients model for grouped data, which he estimates using maximum likelihood estimation.

9. Formally, $\delta' = (\delta_1', \delta_2', ..., \delta_c')$.

10. Amemiya (1978, p. 793) explicitly assumes that all matrices, including Z, are of full-rank. In so doing, he precludes the possibility of multiple group effects.

11. Even if that turns out to be the case, one may find our two-step estimator to be simpler, especially if the group sizes are "unbalanced" and there are multiple groupings in the data.

12. In order to study migration across states, *i* cannot equal *j* for the same observation. Since we limit our study to the 48 contiguous states, we have 48 x 47 = 2256 possible migration flows (M_{ij}'s).

13. Another way is to look at the number of elderly individuals who moved into a location (in-migration) or moved out of a location (out-migration) or the net gain of a location (net in-migration). However, these studies can not account for where migrants moved from or to. Numerous studies focus on the impact of amenities and the cost-of-living on migration, for examples, see Graves and Waldman (1991, county in-migration) and Cebula (1993, state net in-migration). Studies that focus on the public sector consist of Cebula and Kohn (1975, state out-migration), Cebula (1990, state in-migration), Clark and Hunter (1992, county in-migration), and Conway and Houtenville (forthcoming, state in-migration and out-migration). The third way uses individual level data to estimate the probability of moving - Newbold (1996), Kallan (1993) and Dresher (1994), the latter of which considers the public sector.

14. Studies estimating the determinants of migration flows of the *total population* are more numerous and include Fields (1979), Mueser (1989), Frees (1992, 1993), and Gabriel, Mattey and Wascher (1995). Several of these include fixed state effects (via origin and destination dummy variables), either because they have panel data or because they are not interested in group-invariant factors. None to our knowledge have modeled random state effects.

15. In fact, the systematic omission of each state invalidates the usual "within" and "between" transformations typically used for random effects models. A proof is available upon request.

16. This occurs for 139 out of 2256 state combinations. Dropping these 139 combinations from the sample did not significantly alter the results.

17. Note that this effect is not completely captured by distance; Californians may know more about Oregon than New Hampshire residents know about New Jersey and yet the former are a greater distance apart.

18. See Greenwood (1975), Cebula (1979) and Fields (1979) for more discussion of this matter.

Table 1. Definitions, Means, and Standard Deviations
(standard deviations in parentheses)

Dependent Variable		
ln(Flow)	Natural log of the number of individuals aged 65 and over migrating from state *i* to state *j* between 1985-90.[a]	4.485 (2.02)
Explanatory Variables		
Miscellaneous Flow Characteristics		
Distance	The distance between the geographic center of state *i* to state *j* "as the crow flies."	1,034.30 (586.60)
Border	Equals one if state *i* and state *j* border one another, zero otherwise.	0.10 (0.29)
Miscellaneous State Characteristics (state averages)[b]		
ln(Pop)	Natural log of the total state population in 1984.	14.95 (0.97)
Cost of Living	Cost of living index created by McMahon (1991) for 1984. The United States average is normalized to 100.	99.19 (5.68)
Household Income	Median income of households for 1984.	$22,379.66 (1,203.19)
Crime	Total offenses known to police per 100,000 resident population in 1984.	4,547.19 (1,203.19)
Sun	Average percentage of possible sunshine for selected cities (states with more than one city were averaged).	60.20 (7.67)
Heating	Average normal seasonal heating degree days, for periods through 1984. Variable is used to estimate heating requirements.	5,149.85 (2,057.04)
Cooling	Average normal seasonal cooling degree days, for periods through 1984. Variable is used to estimate cooling requirements.	1,162.28 (822.53)
Measures of Publicly Provided Goods[c]		
Education	Per capita general, direct state and local spending on education in 1984.	$727.78 (159.94)
Hospital	Per capita general, direct state and local spending on health and hospitals in 1984.	$184.52 (62.40)
Welfare	Per capita general, direct state and local spending on public welfare, minus Medicaid spending on elderly recipients, in 1984.	$109.48 (48.69)
Medicaid	Total Medicaid spending on elderly recipients per elderly individual in 1984.[d]	

Table 1. Continued

Explanatory Variables		
Measures of the Tax Shares[e]		
Property	Proportion of total general direct state and local spending (G) financed with state and local property taxes in 1984.	0.22 (0.08)
Sales	Proportion of G financed with state and local sales taxes in 1984.	0.17 (0.08)
Income	Proportion of G financed with state and local income taxes in 1984.	0.13 (0.08)
Interact	"Income" times the amount of pension income that is exempt from state personal income taxation in 1984.[f]	445.91 (1,071.73)
All Other T	Proportion of G financed with all other taxes and fees in 1984.	0.52 (0.11)

[a]Taken from *County-to-County migration Flow Files* for the 1990 Census.

[b]Unless otherwise stated these variables are from various editions of the *Statistical Abstract of the United States*.

[c]Unless otherwise stated the source is *Significant Features of Fiscal Federalism*, 1985-86, Table 15.

[d]Since Arizona was exempt from the Medicaid program, we used figures from the Arizona Health Care Containment System. To proxy Medicaid generosity we multiply the total expenditures of Arizona's health care program (which includes recipients of all ages) by the proportion of total United States Medicaid spending on the elderly, then divided this by Arizona's elderly population. This proxies Arizona's Medicaid-type spending on the elderly, per elderly resident. Health Care Financing: Medicare and medicaid Data Book, 1988.

[e]Unless otherwise stated, the source is *Significant Features of Fiscal Federalism*, 1985-86, Table 33.

[f]The source for pension income information is *Significant Features of Fiscal Federalism*, 1985-86, Table 101.

Table 2. Migration Flows
(t-statistics are in parentheses)

Variable		Without Public Sector			With Public Sector		
		OLS	**Random**	**GLS**[b]	**OLS**	**Random**	**GLS**[c]
Constant[a]		-22.964***	2.9302***	2.9302***	-18.328***	2.9302***	2.9302***
		(-22.39)	(15.65)	(15.65)	(-13.98)	(-13.98)	(15.65)
Distance[a]		-0.00104***	-0.00143***	-0.00143***	-0.00120***	-0.00143***	-0.00143***
		(-22.43)	(-33.91)	(-33.91)	(-26.99)	(-33.91)	(-33.91)
Border[a]		1.5818***	1.3259***	1.3259***	1.4767***	1.3259***	1.3259***
		(18.65)	(18.65)	(18.65)	(18.70)	(18.65)	(18.65)
ln(Pop)	o	0.9228***	0.8992***	0.8976***	0.8534***	0.8422***	0.8407***
		(27.21)	(12.55)	(12.52)	(18.40)	(11.51)	(11.49)
	d	0.8761***	0.8525***	0.8509***	0.7209***	0.7096***	0.7081***
		(25.84)	(9.92)	(9.90)	(15.54)	(6.82)	(6.81)
Cost-of-Living	o	0.0089	0.0066	0.0065	-0.0204	-0.0239*	-0.0243*
		(1.21)	(0.42)	(0.41)	(-2.27)	(-1.169)	(-1.72)
	d	-0.0590***	-0.0614***	-0.0615***	-0.0693***	-0.0728***	-0.0733***
		(-8.00)	(-3.27)	(-3.28)	(-7.71)	(-3.62)	(-3.65)
Household Income	o	1.27E-05	1.37E-05	1.38E-05	-2.79E-05**	-2.79E-05	-2.79E-05
		(1.07)	(0.55)	(0.55)	(-2.12)	(-1.34)	(-1.34)
	d	4.65E-05***	4.75E-05	-9.66E-05	-9.66E-06	-9.61E-06	-9.61E-06
		(3.92)	(1.57)	(1.58)	(-0.73)	(-0.33)	(-0.33)
Crime	o	0.000256***	0.000286***	0.000288***	0.000246***	0.000258***	0.000260***
		(9.94)	(5.30)	(5.33)	(8.96)	(5.98)	(6.02)
	d	0.000302***	0.000332***	0.000334***	0.000373***	0.000385***	0.000387***
		(11.71)	(5.12)	(5.16)	(13.58)	(6.27)	(6.30)

Table 2. Continued

Variable		Without Public Sector			With Public Sector		
		OLS	Random	GLS[b]	OLS	Random	GLS[c]
Sun	o	0.02982*** (7.36)	0.03142*** (3.66)	0.03153*** (3.67)	0.01800*** (4.11)	0.01865*** (2.70)	0.01874*** (2.71)
	d	0.03385*** (8.35)	0.03545*** (3.44)	0.03555*** (3.45)	0.02130*** (4.87)	0.02193** (2.23)	0.02201** (2.24)
Heating	o	0.000004 (0.18)	-0.000023 (-0.45)	-0.000024 (-0.48)	-0.000093*** (-3.25)	-0.000115*** (-2.57)	-0.000117*** (-2.63)
	d	-0.000145*** (-6.07)	-0.000172*** (-2.86)	-0.000174*** (-2.89)	-0.000164*** (-5.75)	-0.000186*** (-2.94)	-0.000189*** (-2.99)
Cooling	o	-0.000147** (-2.41)	-0.000230* (-1.81)	-0.000236* (-1.85)	-0.000179*** (-2.85)	-0.000228** (-2.34)	-0.000235** (-2.40)
	d	-0.000310*** (-5.08)	-0.000394*** (-2.58)	-0.000399*** (-2.62)	-0.000266*** (-4.24)	-0.000315** (-2.28)	-0.000322** (-2.32)
Education	o				-0.00186*** (-6.35)	-0.00191*** (-4.15)	-0.00192*** (-4.17)
	d				-0.00177*** (-6.06)	-0.00183*** (-2.79)	-0.00184*** (-2.80)
Hospital	o				-0.00304*** (-7.07)	-0.00323*** (-4.78)	-0.00325*** (-4.81)
	d				-0.00286*** (-6.67)	-0.00305*** (-3.18)	-0.00308*** (-3.20)
Welfare	o				0.00021 (0.29)	0.00029 (0.26)	0.00031 (0.27)
	d				-0.00188*** (-2.61)	-0.00179 (-1.11)	-0.00177 (-1.10)

Table 2. Continued

Variable		Without Public Sector			With Public Sector		
		OLS	Random	GLS[b]	OLS	Random	GLS[c]
Medicaid	o				-0.00026 (-1.80)	-0.00024 (-1.05)	-0.00024 (-1.03)
	d				-0.00066*** (-4.52)	-0.00064* (-1.94)	-0.00063* (-1.93)
All Other G	o				0.00138*** (7.03)	0.00145*** (4.69)	0.00146*** (4.72)
	d				0.00094*** (4.77)	0.00100** (2.28)	0.00101** (2.30)
Property	o				7.457*** (10.28)	7.668*** (6.71)	7.695*** (6.73)
	d				7.564*** (10.43)	7.783*** (4.79)	7.811*** (4.80)
Sales	o				4.091*** (7.12)	4.218*** (4.65)	4.234*** (4.67)
	d				4.983*** (8.67)	5.115*** (3.97)	5.132*** (3.98)
Income	o				3.117*** (4.88)	3.145*** (3.12)	3.148*** (3.12)
	d				5.116*** (8.01)	5.152*** (3.59)	5.157*** (3.59)
Interact	o				-2.97E-06 (-0.10)	-3.82E-06 (-0.08)	-3.93E-06 (-0.09)
	d				5.19E-06 (0.18)	4.04E-06 (0.06)	3.89E-06 (0.06)

Table 2. Continued

Variable		Without Public Sector			With Public Sector		
		OLS	Random	GLS[b]	OLS	Random	GLS[c]
All Other T	o				2.663***	2.684***	2.687***
					(5.48)	(3.50)	(3.51)
	d				3.062***	3.092***	3.096***
					(6.31)	(2.84)	(2.84)
Adjusted R^2		72.8	---	---	76.8	---	---
σ_o^2		---	0.091	0.091	---	0.035	0.035
σ_d^2		---	0.137	0.137	---	0.086	0.086
σ_{od}		---	assumed 0	0.083	---	assumed 0	0.044
ρ_{od}		---	---	0.743	---	---	0.797
LM: $\sigma_o^2 = 0$		---	1,700.99***	1,700.99***	---	1,073.87***	1,073.87***
LM: $\sigma_d^2 = 0$		---	568.83***	568.83***	---	85.98***	85.98***
LM: $\rho_{od} = 0$		---	---	26.47***	---	---	30.52***
LM: $\rho_{od} = -1$		---	---	145.77***	---	---	155.08***

[a]These estimates are from the first stage; see text.

[b]When these estimates are adjusted for estimation error from the first stage $\sigma_o^2 = 0.07$, $\sigma_d^2 = 0.12$, $\sigma_{od} = 0.08$, $\rho_{od} = 0.86$.

[c]When these estimates are adjusted for estimation error from the first stage $\sigma_o^2 = 0.18$, $\sigma_d^2 = 0.07$, $\sigma_{od} = 0.04$, $\rho_{od} = 1.19$.

*, **, and *** refer to 10 percent, 5 percent, and 1 percent significance level, respectively.

References

Amemiya, T. 1978. "A Note on a Random Coefficients Model," *International Economic Review*, 19 (October): 793-96.

Anselin, L. 1988. *Spatial Econometrics: Methods and Models*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Borjas, G.J. and G.T. Sueyoshi. 1994. "A Two-Stage Estimator for Probit Models with Structural Group Effects," *Journal of Econometrics,* 64 (September/October): 165-82.

Bolduc, D., R. Laferriére, and G. Santarossa. 1992. "Spatial Autoregressive Error Components in Travel Flow Models," *Regional Science and Urban Economics,* 22 (September): 371-385.

Case, A.C. 1991. "Spatial Patterns in Household Demand," *Econometrica,* 59 (July), 953-965.

Cebula, R.J. 1979. "A Survey of the Literature on the Migration-Impact of State and Local Government Policies," *Public Finance,* 34: 69-83.

Cebula, R.J. 1990. "A Brief Empirical Note on the Tiebout Hypothesis and State Income Tax Policies," *Public Choice,* 67 (October): 87-89.

Cebula, R.J. 1993. "The Impact of Living Costs on Geographic Mobility," *The Quarterly Review of Economics and Finance,* 33 (Spring): 101-105.

Cebula, R.J. and R.M. Kohn. 1975. "Public Policies and Migration Patterns in the United States," *Public Finance,* 30: 186-196.

Clark, D.E. and W.J. Hunter. 1992. "The Impact of Economic Opportunity, Amenities and Fiscal Factors in Age-Specific Migration Rates," *Journal of Regional Science,* 32: 349-365.

Conway, K.S. and A.J. Houtenville. Forthcoming. "Do the Elderly 'Vote with their Feet'?" *Public Choice*.

Conway, K.S. and A.J. Houtenville. 1998. "Elderly Migration and State Fiscal Policy: Evidence from the 1990 Census Migration Flows," Aging Studies Program Paper No. 13, Center for Policy Research, The Maxwell School. Syracuse, NY: Syracuse University, August.

Dresher, K.A. 1994. "Local Public Finance and the Residential Location Decisions of the Elderly," Ph.D. Dissertation, University of Wisconsin.

Fields, G.S. 1979. "Place-to-Place Migration: Some New Evidence," *Review of Economics and Statistics*, 61 (February): 21-32.

Fournier, G.M., D.W. Rasmussen, and W.J. Serow. 1988. “Elderly Migration as a Response to Economic Incentives,” *Social Science Quarterly,* 69 (June): 245-260.

Frees, E.W. 1992. “Forecasting State-to-State Migration Rates,” *Journal of Business and Economic Statistics,* 10 (April): 153-67.

Frees, E.W. 1993. “Short-term Forecasting of Internal Migration,” *Environment and Planning,* 25: 1593-1606.

Gabriel, S.A., J.P. Mattey, and W.L. Wascher. 1995. “The Demise of California Reconsidered: Interstate Migration over the Economic Cycle,” *Federal Reserve Bank of San Francisco Economic Review,* 0: 30-45.

Graves, P.E. and T.A. Knapp. 1988. “Mobility Behavior of the Elderly,” *Journal of Urban Economics,* 24 (July): 1-8.

Graves, P.E. and P.D. Linneman. 1979. “Household Migration: Theoretical and Empirical Results,” *Journal of Urban Economics,* 6 (July): 383-404.

Graves, P.E. and D.M. Waldman. 1991. “Multimarket Amenity Compensation and the Behavior of the Elderly,” *American Economic Review,* 81 (December): 1374-1381.

Greene, W.H. 1997. *Econometric Analysis.* New York, NY: Macmillan Publishing Company, third edition.

Greenwood, M.J. 1975. “Research on Internal Migration in the United States: A Survey,” *Journal of Economic Literature,* 13 (June): 397-433.

Judge, G.G., W.E. Griffiths, R.C. Hill, H. Lutkepohl, and T. Lee. 1985. *The Theory and Practice of Econometrics.* New York, NY: John Wiley & Sons, Inc., second edition.

Kallan, J.E. 1993. “A Multilevel Analysis of Elderly Migration,” *Social Science Quarterly,* 74 (June): 403-416

McLeod, K.D., J.R. Parker, W.J. Serow, and N.W. Rives, Jr. 1984. “Determinants of State-to-State Flows of Elderly Migrants,” *Research on Aging,* 6 (September): 372-83.

McMahon, W.W. 1991. “Geographic Cost of Living Differences: An Update,” *AREUEA Journal,* 19 (Fall): 426-450.

Moulton, B.R. 1986. “Random Group Effects and the Precision of Regression Estimates,” *Journal of Econometrics,* 32 (August): 385-397.

Mueser, P. 1989. “The Spatial Structure of Migration: an Analysis of Flows Between States in the USA over Three Decades,” *Regional Studies,* 23 (June): 185-200.

Newbold, K.B. 1996. "Determinants of Elderly Interstate Migration in the United States, 1985-1990," *Research on Aging*, 18 (December): 451-76.

Serow, W.J., D.A. Charity, G.M. Fournier, and D.W. Rasmussen. 1986. "Cost of Living Differentials and Elderly Interstate Migration," *Research on Aging*, 8 (June): 317-327.

Voss, P.R., R.J. Gunderson, and R. Manchin. 1988. "Death Taxes and Elderly Interstate Migration," *Research on Aging*, 10 (September): 420-450.

Center for Policy Research
Tim Smeeding, Director
Richard V. Burkhauser, Associate Director for Aging Studies

Center for Demography and Economics of Aging
Douglas A. Wolf, Director

The Maxwell Center for Demography and Economics of Aging was established in September 1994 with core funding from the National Institute on Aging (Grant No. P20-AG12837). The goals of the Center are to coordinate and support research activities in aging, publish findings and policy analyses, and convene workshops, seminars, and conferences intended to train data users and disseminate research findings and their implications to the research and policy communities.

The Center produces two paper series, the Aging Studies Paper Series and Papers in Microsimulation.

Many Center publications are available via Adobe Acrobat format. To access these resources attach to our World-Wide Web home page at http://www-cpr.maxwell.syr.edu.

For more information contact: Martha Bonney (mbonney@maxwell.syr.edu).

Center for Policy Research
426 Eggers Hall
Syracuse University
Syracuse, New York 13244-1020

(315) 443 3114

FAX: (315) 443 1081

AGING STUDIES PROGRAM PAPERS

Paper No.	Title	Author	Date
1	The Division of Family Labor: Care for Elderly Parents	Wolf, Freedman, and Soldo	August 1995
2	Time? Money? Both? The Allocation of Resources to Older Parents	Couch, Daly, and Wolf	December 1995
3	Coresidence with an Older Mother: The Adult Child's Perspective	Soldo, Wolf, and Freedman	December 1995
4	Determinants and Consequences of Multigenerational Living Arrangements: The Case of Parent-Adult Child Coresidence	Schneider and Wolf	January 1997
5	Do Parents Divide Resources Equally among Children? Evidence from the AHEAD Survey	Dunn and Phillips	January 1997
6	Late Life Job Displacement	Couch	January 1997
7	The Immigrant Sample of the German Socio-Economic Panel	Burkhauser, Kreyenfeld, and Wagner	January 1997
8	Testing the Significance of Income Distribution Changes Over the 1980s Business Cycle: A Cross-National Comparison	Burkhauser, Crews, Daly, and Jenkins	June 1997
9	Changes in Economic Well-Being and Income Distribution in the 1980s: Different Measures, Different Outcomes	Burkhauser and Crews	June 1997
10	How Older People in the United States and Germany Fared in the Growth Years of the 1980s—A Cross-Sectional versus a Longitudinal View	Crews and Burkhauser	August 1997
11	Health, Work, and Economic Well-Being of Older Workers, Aged 51 to 61: A Cross-National Comparison Using the United States HRS and The Netherlands CERRA Data Sets	Burkhauser, Dwyer, Lindeboom, Theeuwes, and Woittiez	August 1997

AGING STUDIES PROGRAM PAPERS

Paper No.	Title	Author	Date
12	Estimating Federal Income Tax Burdens for Panel Study of Income Dynamics (PSID) Families Using the National Bureau of Economic Research TAXSIM Model	Butrica and Burkhauser	December 1997
13	Elderly Migration and State Fiscal Policy: Evidence from the 1990 Census Migration Flows	Conway and Houtenville	August 1998
14	Intergenerational Co-Residence and Children's Incomes	Dunn and Phillips	October 1998
15	When Random Group Effects are Cross-Correlated: An Application to Elderly Migration Flow Models	Conway and Houtenville	October 1998