

SOME IMPLICATIONS OF COLLEGE STUDENTS AND MILITARY PERSONNEL FOR THE AGE STRUCTURE OF MIGRATION

*James K. Kindahl and Robert A. Nakosteen**

Abstract—This paper examines the impact of a college or military installation in a metropolitan area on the age structure of the migration for the area. We find that the presence of such institutions causes significant and systematic impacts on the age patterns of in- and out-migration that may not be obvious a priori. Failure to account for these effects in studies of gross migration (whether or not aggregated over age groups), or of net migration by age, will typically lead to bias or to inefficiency in the results, depending on the techniques of analysis used.

I. INTRODUCTION

Economists, regional scientists, and scholars based in other disciplines have long been interested in migration as a means of adjustment to differential economic change among regions within a country. A large number of empirical studies relating to this question have appeared in the literature. Many of these use cross-sectional data from the U.S. Census Bureau. Such studies necessarily incorporate the census definitions and procedures. Sometimes two apparently independent decisions about definitions and procedures affecting the census data can combine to produce difficulties in use not obvious from a consideration of either one separately. The existence of these problems can easily be overlooked. One such problem occurs in the calculation of migration rates for areas in which college students and/or military personnel comprise a substantial proportion of the measured population of an area. A significant proportion of the Metropolitan Statistical Areas (MSAs) in 1980 had this characteristic (Long 1983).

The problem arises from the conjuncture of, on the one hand, the census definition that includes both college students and military personnel as residents of the census area in which they live as of April 1 of the census year; and on the other hand, of the measurement of migration as change of residence between the census date five years prior to the census year and the census year itself (the "five-year migration" definition). A large proportion of people who were college students five years prior to the census year would not be college students in the census year; similarly, a large proportion of people who were college students in

University of Massachusetts/Amherst.

the census year would not have been college students five years previously. Those who were in college in the census year include a substantial proportion who changed MSAs of (census-defined) residence to do so, and hence are considered in-migrants to the MSA of their college. Those who were in college five years previously, but no longer, and who did not remain in the college MSA after leaving college are considered out-migrants from the MSA of their college. Consequently, MSAs hosting a large number of college students in proportion to their total population will be shown as having significantly larger in-migration and out-migration rates than other areas, *ceteris paribus*, even when age is ignored. A similar pattern will be found for MSAs in which military personnel form a substantial proportion of the population because of the presence of a large military base. Here the result arises because of the high rates of movement of individuals into and out of the military over a five-year period, as well as the high transfer rates of military personnel among bases.

In terms of the five-year age intervals typically used in census reports, the effect of both college-student and military movement in the MSAs strongly affected by these movements is to increase the measured rates of gross in-migration in the age interval 15-19 years and to increase the measured rates of out-migration in the age interval 25-29 years (a result, at least for the 25-29 age interval, which seems to the authors to be counter-intuitive). The reasons for this expectation are shown in section III. Note that the net migration rates for those two age categories will be affected as well. Net migration in the 15-19 age group will tend to be higher, and net migration in the 25-29 age group will be lower (or more negative because we will define net migration as net in-migration, or in-migration minus out-migration). If the data are aggregated over age groups, the net migration will be unaffected by this phenomenon, but the gross in-migration and the gross out-migration rates will be affected.

Thus, measured net and gross migration rates in some age categories—and therefore the gross migration rates for the entire population—will be significantly affected in certain MSAs by the presence of colleges and military bases. However, other MSAs will be largely unaffected, either because there are no colleges or military bases in the MSA, or because the population of the MSA is large relative to the population of college students and military personnel. It will be the classic "college town," whose main economic base is centered around a university or a college, or its military analogue, in which the effect will be strongest. For studies attempting to examine the role of migration as an adjustment mechanism, this effect can be a serious problem. These are not random fluctuations, but systematic factors affecting certain migration rates for certain identifiable MSAs consistently; they are factors that are largely unrelated to migration as an adjustment mechanism. Further, they occur in precisely the age groups in which adjust-

ment to economic conditions normally takes place. Thus, use of the entire set of MSA data in a cross-sectional study could produce misleading results unless some means of dealing with this problem were employed. Suppose that regression analysis were used, for example, to explain the rates of gross or net migration for individual age groups, using data on economic conditions by MSA as explanatory variables. Failure to take college student movement into account would surely lead to a substantial reduction in explanatory power because of the absence of an important variable determining the dependent variable. If the omitted variable (the proportion of transient college students and military personnel in the population) were correlated with one or more of the included independent variables, the result would be bias and inconsistency in the estimated coefficients—a serious problem indeed. Similarly, an analysis of tabular data could suffer because particular cells would be strongly affected by factors not accounted for in the analysis. At best, failure to take the college/military phenomenon into account would lead to low explanatory power and confusion; at worst, it would give totally misleading results (as it would if included variables were highly correlated with the underlying omitted one).

All migration studies using census data will be affected, to a greater or lesser degree, by this problem, except for two general classes of studies: first, studies based on microdata and second, studies based on aggregated data which use population data aggregated over all ages and which use net migration measures only, not measures of gross migration.

Microdata studies based on files of data in which the basic unit is the record of an individual or a family have been a prominent focus of attention over the past decade or so (Greenwood 1985). In these studies, it is possible to identify and isolate those individuals for whom migration is related to either college attendance or military service. Eliminating such individuals from the study avoids the problems created by college and military migration, and this typically is done.

Studies in which data are aggregated over all ages and which use net migration only (i.e., no gross in- or out-migration figures) are also immune to the problem. Examples of such studies are Lowry (1966) and Sommers and Suits (1973). Because the age structure is suppressed, variations in migration rates by age are not recorded, and because net migration is used, the increased in-migration is canceled by the increased out-migration. In fact, this type of study is the only one in which no impact from college or military migration is recorded.

All other migration studies will be affected. Thus any aggregated (i.e., non-microdata) study which uses either an age breakdown or gross migration (e.g., Long [1988] and Wilson [1988]) will reflect the impact of the census conventions. Aggregate studies seem to be making a comeback after being largely replaced in

the literature by microdata studies for a number of years. Many, but not all, aggregate studies account for college and military migration.

To aid in understanding the impact of college and military migration on migration studies, consider a cross-sectional study of the relationship between economic conditions and gross out-migration for MSAs. Assume that there are two MSAs in the study, identical except that one has a large college student population and the other does not. Then the same economic conditions will be associated with greater out-migration in the college MSA than in the other MSA, diluting the effect of economic conditions. A more damaging situation would occur if the college MSA exhibited economic conditions better than those in the other MSA. There would be a good chance that the college MSA would have higher gross out-migration than the non-college MSA, making it appear that better economic conditions lead to more out-migration, not to less out-migration, as one would expect.

II. THE ANATOMY OF COLLEGE STUDENT/MILITARY MIGRATION

Census migration data are based upon the five-year retrospective residence question. Individuals are considered to have migrated from Area A (possibly an MSA) to Area B (possibly another MSA) if the area of current residence differs from the area of residence five years previously. In data based on this convention, MSAs with a relatively large proportion of the population as college students and/or military personnel unambiguously exhibit large rates of in-migration for the group aged 15-19 years at the end of the migration interval and, also unambiguously, exhibit large rates of out-migration for the group aged 25-29 years. The group aged 20-24 years—which one might normally associate with out-migration from college towns—actually exhibits a mix of in- and out-migration related to college/military attachment. The cause of the somewhat counter-intuitive result requires some exegesis of the census protocols.

To bring out the mechanics of this result, we refer specifically to data from the 1980 census on gross in- and gross out-migration by age category (in 1980), for the 318 MSAs during the migration interval 1975-80. Consider an MSA that is strongly affected by college student migration. The census definitions classify a person as an in-migrant to this MSA if that person was not a resident in April 1975 and was a resident in April 1980; a person is an out-migrant if the person was a resident in April 1975 and not a resident in April 1980. We carry out the analysis for a student of modal age for migration associated with college entrance (assumed to be 18), who enters college in September. The modal student who

migrated to enter college in this MSA at the beginning of a term in September in any of the years 1976 through 1979 and left this MSA four years later would be reported in the census statistics as an in-migrant to this MSA since those students would have finished in May of any of the years 1980 through 1983 and hence would have been present in this MSA in April 1980. In April 1980, students who had entered in September of 1976 would be 21 or 22; those who had entered September of 1977 would be 20 or 21; those who had entered in September of 1978 would be 19 or 20; and those who entered in September of 1979 would be 18 or 19. Thus the impact of college entrance on measured in-migration in this MSA will affect both the census age categories 15-19 and 20-24, because the reported data are all for ages as of April 1980 (the end of the five-year period).

A similar analysis shows the effect on measured out-migration from this MSA of students leaving at the modal age of 22. To affect measured out-migration, a student would have had to be living in this MSA in April 1975, and not living there as of April 1980. Thus the student would have had to enter college not earlier than September 1971 (in order to be still in residence in April 1975), and not later than September 1974 (in order to have been present for the census in March 1975). (Those who entered in September 1970 and earlier would have graduated prior to April 1975 if they had finished in four years.) Those who had entered in September 1971 at age 18 would have been 26 or 27 in April 1980; the other classes would have been 25 or 26, 24 or 25, and 23 or 24, for the classes entering 1972, 1973, and 1974, respectively. Hence the age categories in April 1980 of those who contributed to measured out-migration (but not measured in-migration) for the college town would be both the 20-24 and 25-29 age categories.

Note that the distinguishing characteristic of those who enter the migration statistics is that they were in college in April 1975 or April 1980, but not both. No student who spent the modal time of four years in college could be in college in both 1975 and 1980. Essentially, those who were in college in 1975 contribute to measured out-migration, those who were in college in 1980 contribute to measured in-migration. Note also that a student who entered college in September 1975 and left (after graduation) in May 1979 would not be registered as either an in-migrant or an out-migrant.¹

Migration in MSAs in which military installations are important is less clear because there is so much migration between military bases, much of it not related directly to age. Nevertheless, there are clearly movements in and out of the military as a whole: most people who join the armed forces do so between the ages between 17-24 and therefore create a strong gross in-migration to military areas in the age groups 15-19 and 20-24. The time of leaving the military is much more variable. Nevertheless, a substantial fraction leave after an initial three- or four-year term and hence, there is substantial gross out-migration from the

military during the ages 20-28. Putting the two together, it is clear that the effect will be to increase in-migration into MSAs in which military bases are located in the age group 15-19, and out-migration in the age group 25-29. The effect on the 20-24 age group is once again, ambiguous. To the extent that individual military bases mirror the net movement into and out of the military establishment, this will be the effect on the corresponding MSA migration rates. Not all individual bases may do so, but military bases in general should. Hence we should expect that some MSAs that include military bases will have the characteristics of college towns in this respect, though there may be MSAs that include military bases which do not have those characteristics.

III. IMPLICATIONS FOR EMPIRICAL DATA

Not all students who attend college are ipso facto migrants in the census definitions. Some students live at home and commute to college; others live on or near the campus, but remain in the same MSA because they attend a college in the MSA of their former home.² The analysis applies to students who attend college in an MSA (or state) other than their immediately prior home. We call the first group (those who do migrate to attend college) *transient* students, and the second group (those who do not migrate to attend college) *local* students.

As a first approximation, we assume that the ratio of local students to the total population, or to the population in the student age groups, does not vary systematically over MSAs; in particular, we assume that this ratio does not vary with either the population of the MSA or with the number of college students in the MSA. As a first approximation, this is a reasonable assumption: on the supply side, there is no reason to expect that the ratio of the number of spaces in local colleges available for residents of the MSA should either rise or fall with the size of the MSA; and on the demand side, there is no obvious reason why the demand per resident for spaces in local colleges would vary with the extent to which education is an important industry within the MSA.³

These assumptions imply that the expected number of local students in each MSA will be a constant fraction of the local population; the number is assumed to be determined essentially by the demand for a college education. Surely, though, the number of transient students is not determined by the size of the local population. We assume that the number of transient students is, in the short run, determined largely by the capacity and character of the colleges located within the MSA—a capacity which can be treated as exogenous at any point in time and which surely bears no strong systematic relation to the size of the population. By

contrast, the assumptions imply that the number of places for local students, in city and state colleges especially, would vary directly with population size.

In a cross section, one component of student population (local students) should be proportional to MSA population size, and the other (additive) component (transient students) should be exogenous and not related to MSA population size.⁴ This assumption is consistent with our sense that, as MSAs grow, the number of transient students studying in the MSA does not grow in consequence; that is, population growth of an MSA would stimulate the growth of colleges serving the local market, but does not of itself stimulate the growth of colleges serving transient students. We take this as a given. Hence there is an upper limit to the absolute number of transient students in any one MSA, but no limit to the absolute number of local students. It follows that in cross-sectional data, as the total population of an MSA increases, the proportion of transient students falls, on the average, after some population size, even though the proportion of local students remains constant. Consequently, the proportion of students in a population falls as the population increases, at least after some population size is reached.

Our expectation is that some MSAs will exhibit migration patterns that are distinguished from other MSAs in that there will be significantly higher in-migration rates in the 15-19 age group, and significant higher out-migration rates in the 25-29 age group, than in other MSAs, as a result of the in- and out-migration patterns of college students and military personnel in these MSAs. Those are the MSAs that we have dubbed "college towns."⁵ We also expect that the MSAs that exhibit this pattern will contain smaller populations than those which do not.

IV. EVIDENCE

To test these expectations, we have calculated the net migration rates in those two age categories for all 318 MSAs in the 1980 U.S. Census. The net migration rate is defined as the gross in-migration minus the gross out-migration, divided by the population in the age group at the end of the migration interval. We have then normalized each migration rate by subtracting from each rate the mean rate in the same age group for all MSAs, then dividing the result by the standard deviation of the latter. The normalized net migration rate thus expresses each MSAs rate as the number of standard deviations from the mean rate for all MSAs.

As summary statistics, we have calculated the algebraic difference between the normalized net migration of 15-19-year-olds and the normalized net migration of 25-29-year-olds. For convenience, we call this statistic the "net migration difference." This difference will be positive and large for those areas that exhibit the pattern we expect from MSAs which are college and military towns.

A listing of the 50 MSAs with the highest values of the net migration difference comprises the first part of Table 1; data are arranged in descending order of net migration difference. For each MSA, the value of the net migration difference and the values of the difference and the normalized net migration rates for the two relevant age groups are given. The names of large colleges or military bases in or near the MSA are given also. If our expectation is correct, we should find that most MSAs on this list have colleges or military bases in their midst and that their aggregate student and/or military population is large, relative to the total population of the MSA. This is, in fact, what we do find. The first 50 MSAs on our list contain colleges or military bases that are clearly important components of the economy of the MSA. Among the first five on the list are several well-known college towns: State College, Pennsylvania; Lawrence, Kansas; Bloomington, Indiana; and College Station, Texas (part of the Bryant-College Station MSA). Leading the list is a relatively small MSA with a large military base: Jacksonville, North Carolina, home of Camp Lejeune.

The result of this impressionistic categorization of MSAs by influence of colleges and military bases is certainly consistent with our hypothesis. A more precise test of our hypothesis begins with the implication of our analysis: the net migration difference should be greater, the higher the proportion of transient college students and military personnel in the population. A good test of this implication therefore would be to relate the proportion of those groups in the population of each MSA to the net migration difference of the same MSA. This test cannot be carried out because there is no way to separate transient from local students. As a substitute, we use the proportion of all college students in the population in lieu of the proportion of transient students in the population. Because the proportion of transient students in the college student population will tend to decline, on the average, with population size, the analysis is less straightforward than it otherwise would be. In particular, although we would expect that the net migration ratio would rise continually as the proportion of transient students in the population rises, that relation would not necessarily hold unequivocally when the proportion of all college students in the population is substituted: the relation of the proportion of college students in the population to the net migration ratio will become weaker, the larger the population size. We adapted our test procedures to the exigencies of the data.

We defined the group composed of the 50 (out of 318) MSAs with the highest net migration difference as college towns. Our predictions then are that (1) the mean net migration difference of this group should be significantly greater than that of the remaining 268 MSAs, and (2) the mean population size of the first 50 should be significantly lower than that of the remaining 268. The tests were run on that basis.

TABLE 1
Net Migration for Age Groups 15-19 And 25-29 and Net Migration Difference for Metropolitan Statistical Areas
(Ranked by Net Migration Difference. Includes the Highest 50 and Every 20th MSA Thereafter.)

	Rank	Dif- ference	15-19	25-29	Ratio
JACKSONVILLE, NC (Camp Lejeune)	1	8.379	2.786	-5.593	0.336
STATE COLLEGE, PA (Penn. State Univ.)	2	7.699	3.217	-4.482	0.252
LAWRENCE, KS (Univ. of Kansas)	3	7.525	3.293	-4.232	0.267
BLOOMINGTON, IN (Univ. of Indiana)	4	7.318	3.445	-3.873	0.257
BRYAN-COLLEGE STATION, TX (Texas A&M Univ.)	5	7.278	4.303	-2.975	0.319
CHAMPAIGN-URBANA-RANTOUL, IL (Univ. of Illinois)	6	7.031	3.606	-3.425	0.199
LAFAYETTE-WEST LAFAYETTE, IN (Purdue Univ.)	7	6.714	3.294	-3.421	0.216
COLUMBIA, MO (Univ. of Missouri)	8	6.138	3.061	-3.077	0.156
GAINESVILLE, FL (Univ. of Florida)	9	5.146	2.871	-2.274	0.179
BLOOMINGTON-NORMAL, IL (Illinois State Univ.)	10	5.007	2.530	-2.477	0.151
PROVO-OREM, UT (Brigham Young Univ.)	11	4.513	2.494	-2.019	0.110
ATHENS, GA (Univ. of Georgia)	12	3.990	2.099	-1.892	0.143
KILLEEN-TEMPLE, TX (Fort Hood)	13	3.963	1.883	-2.080	0.215
TALLAHASSEE, FL (Florida State Univ.)	14	3.903	2.311	-1.593	0.102
IOWA CITY, IA (Univ. of Iowa)	15	3.867	2.031	-1.836	0.215
MUNCIE, IN (Ball State Univ.)	16	3.645	1.049	-2.596	0.112
ANN ARBOR, MI (Univ. of Michigan)	17	3.291	2.015	-1.276	0.082
TUSCALOOSA, AL (Univ. of Alabama)	18	3.246	1.775	-1.471	0.093
LA CROSSE, WI (Univ. of Wisconsin/La Crosse)	19	3.186	1.993	-1.193	0.089
LAWTON, OK (Fort Sill)	20	3.175	2.246	-0.929	0.178
GRAND FORKS, ND-MN (Univ. of N. Dakota; Grand Forks AFB)	21	3.106	1.008	-2.098	0.137
LINCOLN, NB (Univ. of Nebraska)	22	3.090	1.529	-1.561	0.086
BANGOR, ME (Univ. of Maine/Orono)	23	2.958	1.551	-1.407	0.158
LUBBOCK, TX (Texas Tech. Univ.)	24	2.619	0.993	-1.626	0.098
WACO, TX (Baylor Univ.)	25	2.609	1.444	-1.165	0.058
CHARLOTTESVILLE, VA (Univ. of Virginia)	26	2.569	1.902	-0.667	0.102
CHICO, CA (California State Univ./Chico)	27	2.311	1.337	-0.975	0.080
FAYETTEVILLE, NC (Fayetteville State Univ.; Ft. Bragg)	28	2.303	0.926	-1.377	0.200
CLARKSVILLE-HOPKINSVILLE, TN-KY (Austin Peay Univ.; Ft. Campbell)	29	2.237	1.456	-0.781	0.160
EAU CLAIRE, WI (Univ. of Wisconsin/Eau Claire)	30	2.204	1.138	-1.067	0.072
FORT COLLINS, CO (Colorado State Univ.)	31	2.185	2.417	0.232	0.097
LAS CRUCES, NM (New Mexico State Univ.)	32	2.100	0.903	-1.197	0.119
AUSTIN, TX (Univ. of Texas)	33	2.045	1.551	-0.493	0.077
SPRINGFIELD, MO (South-West Missouri State College)	34	2.029	1.308	-0.721	0.093
MADISON, WI (Univ. of Wisconsin/Madison)	35	2.023	1.044	-0.980	0.086
FARGO-MOORHEAD, ND-MI (North Dakota State Univ.)	36	1.930	1.274	-0.656	0.055
BURLINGTON, VT (Univ. of Vermont)	37	1.858	1.292	-0.566	0.069
LANSING-EAST LANSING, MI (Michigan State Univ.)	38	1.838	0.830	-1.008	0.065
BELLINGHAM, WA (Western Washington Univ.)	39	1.831	1.227	-0.604	0.575

TABLE 1 (Continued)
 Net Migration for Age Groups 15-19 And 25-29 and Net Migration Difference for Metropolitan Statistical Areas
 (Ranked by Net Migration Difference. Includes the Highest 50 and Every 20th MSA Thereafter.)

	Rank	Dif- ference	15-19	25-29	Ratio
COLORADO SPRINGS, CO (U.S. Air Force Academy)	40	1.803	0.505	-1.298	0.099
ABILENE, TX (Abilene Christian Univ.)	41	1.754	0.810	-0.944	0.061
RALEIGH-DURHAM, NC (Duke Univ., U. of NC, NC State U.)	42	1.575	1.333	-0.243	0.040
BILOXI-GULFPORT, MS (Keesler Air Force Base)	43	1.572	0.769	-0.802	0.073
FAYETTEVILLE-SPRINGDALE, AR (Univ. of Arkansas)	44	1.562	1.531	-0.031	0.057
HAMILTON-MIDDLETOWN, OH (Miami Univ.)	45	1.546	0.969	-0.577	0.067
GREELEY, CO (Northern Colorado Univ.)	46	1.510	1.152	-0.358	0.059
ST. CLOUD, MN (St. Cloud State Univ.)	47	1.409	0.656	-0.753	0.071
SAN ANGELO, TX (Angelo State Univ.; Goodfellow AFB)	48	1.393	1.493	0.100	0.073
WICHITA FALLS, TX (Midwest. State Univ.; Sheppard AFB)	49	1.210	0.714	-0.496	0.099
TERRE HAUTE, IN (Indiana State Univ.)	50	1.183	0.448	-0.735	0.055
SALINAS-SEASIDE-MONTEREY, CA	70	0.452	0.198	-0.254	0.077
ALEXANDRIA, LA	90	0.099	-0.362	-0.461	0.022
SAN ANTONIO, TX	110	-0.157	0.061	0.218	0.060
NASHVILLE-DAVIDSON, TN	130	-0.352	0.231	0.583	0.018
GRAND RAPIDS, MI	150	-0.472	-0.167	0.305	0.008
PHILADELPHIA, PA-NJ	170	-0.601	-0.649	-0.048	0.017
BRIDGEPORT, CT	190	-0.712	-0.932	-0.221	0.006
PATERSON-CLIFTON-PASSAIC, NJ	210	-0.841	-1.076	-0.235	0.000
HAGERSTOWN, MD	230	-0.958	-0.409	0.549	0.010
TEXARKANA, TX-TEXARKANA, AR	250	-1.055	-0.290	0.766	0.001
ST. LOUIS, MO-IL	270	-1.179	-0.996	0.183	0.007
ANCHORAGE, AK	290	-1.331	-1.021	0.309	0.095
SPRINGFIELD, IL	310	-1.701	-1.072	0.629	0.024

Source: Calculated from the U.S. Census Bureau, *Barron's College Guide*, and *Peterson's Guide to Four-Year Colleges*.

Note: The column headed "15-19" is the gross in-migration rate (gross in-migration/population) for the age group 15-19; the column headed "25-29" is the gross out-migration rate for the 25-29 age group; the column headed "Diff" is the net migration difference, defined as the algebraic difference: gross in-migration rate for 15-19 age group minus gross out-migration rate for 25-29 age group; the column headed "Rank" is the rank (out of 318) of the gross migration difference.

The choice of 50 was arbitrary. Too low a number would have led to a low power of the tests and a consequent large risk of not finding significance even though the null hypotheses (of equal means) were false. Too large a number would have destroyed any meaningful interpretation of the characterization of these MSAs as college towns. By taking a number as great as 50, we have erred, if at all, on the side of characterizing too many MSAs as college towns.⁶ We preferred accepting that risk to accepting the risk of taking too few.

The data on college student population used in this study comes from *Barron's College Guide* for 1988 and *Peterson's Guide to Four-Year Colleges* for 1989; these were taken as approximations to the 1980 population. The figures for all four-year colleges are given by state; aggregating the figures for each MSA gave the college population. We used the census estimates of armed services population by MSA as our measure of military base population by MSA. These figures were compiled for each of the first 50 MSAs in order by net migration difference, and every twentieth entry beyond that (i.e., the 70th, 90th–310th in the list totaling 13, which we dubbed the "non-college MSAs"). The resulting sample was 50 college town MSAs and 13 other MSAs, which were a systematic random sample from the remaining group. The sample of the remainder might seem small, but the only danger is that the power of the test would be too low to show significance. Such was not the case, as is shown below; and so there was nothing to be gained in pursuing the rather tedious task of compiling college and military population figures by MSA.

The hypothesis is that the members of the "first 50" MSAs—the ones displaying the highest net migration differences and thereby exhibiting the relations that the analysis implies would be most pronounced in college towns—will have on average higher ratios of college students plus military personnel to total population than the rest of the MSAs represented here by the systematic random sample of 13. Our test statistics are the (unweighted) means of the ratios for each subgroup; the figures are 0.136 and 0.026, respectively.⁷ The test used is the non-parametric Wilcoxon test (equivalent to a Mann-Whitney test), which requires no assumptions about the form of the underlying statistical distribution.

The principle of the Wilcoxon test is simple: if the null hypothesis that the means were the same were true, then when all 63 observations (50 and 13 respectively) are put together, the ranks of the elements of the two groups should be randomly distributed. On the other hand, if the mean of the ranks of the first group is higher (when ranked from the highest to the lowest) than the mean of the ranks of the second group, then that is evidence in favor of the alternative hypothesis that the mean value of the first group is higher than that of the second. Because this is a nonparametric test, no troublesome assumptions about distribution or

equality of variances of the two groups are required. The test is significant at the .0001 level. Clearly the sample size of 13 is sufficient to establish the fact.

The ratio of college students plus military personnel to total population tends to fall as rank increases throughout the listing of the first 50 MSAs, except for minor aberrations; MSAs with a high net migration difference have, for the most part, high ratios of college students plus military personnel to total population. For the first 50 MSAs, the Spearman rank correlation coefficient r_s was $-.747$; for the 13 chosen in the systematic random sample from the remaining 268 MSAs, r_s was $-.231$, and for the combined sample of 63, r_s was $-.827$. Significance tests were performed on all three samples using the normal approximation to the appropriate transformation of r_s . The Spearman coefficients of the two larger samples (both of which include the first 50 MSAs) are both significantly different from zero at the .0001 level (approximate t values of -7.78 and -11.50 , respectively); the sample of 13 from the MSAs not in the first 50 was nowhere near significant (approximate t value of -0.23).

The other hypothesis, that the mean population in college towns would be lower than the mean population in the remaining ones, is also borne out by the evidence. The calculations were carried out using two different measures of population in turn: the total population including college students and military personnel, and the total population minus college students and military personnel in the MSA. Ideally, the measure would be total population minus transient college students and military personnel, but that is impossible. Here the means of the total populations of the two groups were 162,015 for the college towns, and 821,220 for the sample of the remainder, for total population including college students and military personnel, and 141,158 and 805,194 for the two groups, respectively, when college students and military personnel are omitted entirely. Such differences hardly require significance tests. As it turns out, both tests are significant at the .0001 level using the Wilcoxon test.

V. CONCLUSIONS

The empirical evidence indicates clearly that some MSAs have measured migration rates for certain age groups that are strongly affected by the presence of college students and military personnel, a consequence of the census conventions. This systematic increase in gross migration rates for certain MSAs is unrelated to adjustments to changed economic conditions in the MSAs in question. Inclusion of those most heavily involved in this "migration" of college students and military personnel is likely to lower, perhaps substantially, the explanatory power of most any model based on comparisons of economic condi-

tions in two or more MSAs (or other sub-state areal units). If rates of gross out-migration (or gross in-migration) for the total population are chosen as the variables to be explained, the MSAs with the largest rates of migration (whether defined as gross out-migration or gross in-migration) will tend to be those with the highest proportion of college students and military personnel. If the ratio of college students and military personnel to total population is not built into the model, regression analyses using gross in- or gross out-migration as dependent variables will be biased if any of the independent variables are correlated with that ratio (for an independent variable would be correlated with an omitted relevant variable and, therefore, with the error term in the equation). At best, no independent variable in the equation would be correlated with the ratio of college students plus military personnel to total population, in which case there would be an increased variance of the estimated coefficients, induced by the fact that a relevant—and important—variable had been omitted.

If the model were to focus on the age groups most likely to respond to economic change by changes in migration decisions—the young adult age groups—the problem would be accentuated; for these are the age groups most bound up in the college student/military personnel effect on measured migration.

Thus, our findings suggest that any use of the census migration data on the 318 MSAs of 1980 or any comparable data set as raw data for cross-sectional studies should take these relations into account. A typical aim of economic studies, for example, is to relate variation in migration rates to economic variables. Use of the MSA data for such purposes is complicated by the relations enumerated above. If the age detail is to be used, avoidance of bias or gross inefficiency will typically require either incorporation of these relations into the model itself, or exclusion of the MSAs in which migration in the young age groups is dominated by these factors. Even if no age or gender detail is to be used, and only aggregate migration is employed, the problem still exists for gross migration statistics: the college MSAs will have both higher gross in-migration and gross out-migration rates because of college/military migration. Only if net migration is used, with no age/gender breakdown, can the relation be ignored with impunity.

The remarks about cross-sectional studies do not apply to studies based directly on microdata, because the latter are cross-sectional studies of individuals, not of MSAs. Of course, such studies are incapable of providing information on MSAs as such except by aggregation; and should that be done, the problem would crop up once again in the aggregation process. Indeed, the census data are themselves the results of such aggregation. The problem discussed arises whenever the unit of analysis is an aggregate of individuals (such as an MSA) rather than of in-

dividuals alone. Many of the important questions refer to such aggregates, however; these cannot be answered from studies of individuals alone.

ENDNOTES

1. The effect of having some students stay less than four years is to reduce the proportion classified as in-migrants in the 20-24 age category relative to the 15-19 category. The effect of a student staying longer than four years is exactly the opposite of the effect of staying less than four years: the proportion in the 20-24 age class (and possibly higher classes) is raised relative to the proportion in the 15-19 age class.

2. Such an occurrence happens occasionally with migration of military personnel also, but such cases are unusual.

3. If the discussion concerned individual counties, or individual cities and towns, there may well be some effect in that a very small town with a large college may have a larger proportion of educated parents, such as faculty members. These parents may have a substantial effect on the proportion of the population that attended college. The MSA is a large enough unit that this factor should not present a serious problem.

4. The idea that transient student populations bear little relation to MSA populations (at least excluding the students themselves from the MSA population) is certainly consistent with anecdotal evidence. Columbia University, for example, hardly dominates New York City in the way that Pennsylvania State University dominates State College, Pennsylvania. Numerous other examples illustrate the same point.

5. We have elected to use the terms "college towns" and "non-college towns" rather than the formally correct terms "college/military towns" and "non-college/military towns" to make the expressions less cumbersome.

6. The data have been ordered by the degree of conformity to the "college town" predictor (the net migration difference). The larger the number of MSAs included in the "college town" list, the less should be the mean difference between this group and the rest of the MSAs on any relevant statistic, and hence the power of the test should be lower, the greater the number chosen as "college and military towns."

7. Our hypothesis is that the means of the ratios in the two samples of 50 and of 13, respectively, will be higher in the sample of 50 than in the sample of 13. Note that this method of computing means has the consequence that multiplying the means by the number in the sample will not give the ratio for the sample as a whole. The ratios of the sum of all college and military personnel in each

subsample, divided by the sum of all populations in each subsample, are 0.117 and 0.020, respectively; the means of the ratios of each subsample (i.e., the means of 50 and 13 ratios) are 0.136 and 0.026, respectively, as indicated in the text. The first definition computes the weighted mean of the ratios, the weights being the populations of the MSAs. The second computes the unweighted mean of the ratios. The fact that the weighted means are lower than the unweighted means indicates that in the sample, the larger MSAs tend to have lower ratios than do the smaller MSAs. This is, of course, consistent with our hypothesis.

REFERENCES

- Barron's College Guide*. 16th ed. New York: Barron's Educational Series Inc., 1988.
- Greenwood, Michael J. "Human Migration: Theory, Models, and Empirical Studies," *Journal of Regional Science* 25, no. 4 (1985): 521-544.
- Long, John F. "The Effects of College and Military Populations on Models of Interstate Migration," *Socio-Economic Planning Sciences* 17, no. 5-6 (1983): 281-290.
- Long, Larry. *Migration and Residential Mobility in the United States*. New York: The Russell Sage Foundation, 1988.
- Lowry, I.S. *Migration and Metropolitan Growth: Two Analytical Models*. San Francisco: Chandler Publishing Co., 1966.
- Peterson's Guide to Four-Year Colleges*. Princeton: Peterson's Guides, 1989.
- Sommers, P.M., and D.B. Suits. "Analysis of Net Interstate Migration," *Southern Economic Journal* 40, no. 2 (1973): 193-201.
- Wilson, Franklin D. "Components of Change in Migration and Destination-Propensity Rates for Metropolitan and Nonmetropolitan Areas: 1935-1980," *Demography* 25, no. 1 (1988): 129-139.