

The Hidden Component in Census-Derived Migration Data: Assessing Its Size and Distribution

Beryl Nicholson

Comparison of Norwegian "linked" decennial census data with statistics compiled from 10 years' migration registrations showed that the amount of movement omitted by census data was considerable. This hidden movement was of a similar order at every administrative level, but only when total movement was considered. There was wide variation between regions, migration directions, and streams. In some cases census data misrepresented the direction of net movement. Available evidence suggests that these patterns are not confined to Norway, raising the possibility that research findings based on census-derived migration data may merely be artifacts of the data.

Migration data derived from censuses and surveys that obtain information on place of residence at a previous date are known to be less comprehensive in their coverage of movement than those based on records of all moves, which are produced by systems of migration registration. Beyond that, little attention is given to any further deficiencies there might be in census-derived data. Of course, they might be so minor as not to affect the validity of research findings. Yet we do not know whether they are or not, for apart from limited investigations that have indicated the volume of movement that goes unrecorded in the census (e.g., Courgeau, 1979, pp. 1-29; Rees, 1977, pp. 247-272), no substantial research has been done on this question.

The Present State of Knowledge

A certain amount of information about the degree of adequacy of census data can be inferred from the limitations imposed by the method of compilation. We already know that census data omit some types of movement, notably repeated movement, which would be recorded by a system of continuous registration. We can deduce that an age, and probably also a sex, bias will result from the factors that render those moves hidden.

What we do not know is the proportion of movement that is omitted—though indications are that it is considerable (DaVanzo & Morrison, 1981, p. 93)—the degree of variation between one area and one migration stream and another, and the nature and extent of bias in the data.

To investigate these questions, one needs parallel sets of data from the census and registration, and these are rarely available. When registration data exist, there is no reason to include retrospective questions on residence in the census. Only when data similar to those produced by the census have been obtained for other reasons can they be compared with registration data. This article attempts to clarify some of the methodological problems inherent in using census-derived data to measure migration on the basis of one such comparison.

The Data

The sources for the comparison are Norwegian migration statistics, which are compiled annually from migration registrations, and "linked" Norwegian census data. In 1970 and again in 1980, the individual records of each person counted in the decennial census were linked to the record for that same person from the previous census, provided they resided in Norway on both census dates. The linking was effected by matching the personal identification number of each individual. In only a small number of cases, estimated at between 13,000 and 18,000 cases (equivalent to .4-.57% of the population enumerated at both censuses), was there a failure to match records from the 1960 and 1970 censuses, due to errors, especially in the 1960 data (Byfuglien, 1977, p. 161). Increasing technical sophistication probably further reduced the errors in the following decade, therefore the data can be considered to be of high quality.

The published tables for each county show the population by commune of residence (the commune is the smallest administrative unit) within that county at two successive censuses and those who had entered or left the county in the interval (County Statistics 1977, 1983).¹ The change of residence data thus generated are comparable to those obtained from a retrospective census question on residence, though more accurate. There is no breakdown by sex or age.

Migration statistics show the number of moves *registered* in each calendar year. Published tables show moves between administrative and statistical areas for each year, with some breakdowns by sex and age. Total migration is defined as the number of registered moves that cross a commune boundary. Every person, including children, who changes domicile from one commune to another is obliged by law to register the move with the commune authorities within 8 days. Despite strict enforcement, a small proportion of moves goes undetected and not all moves are registered in the year that they occur. The level of hidden movement will therefore be slightly understated when these data are used as the basis of a comparison.

The census is taken on November 1, therefore there is a discrepancy of 2 months in the time periods covered by census data and 10 years' migration statistics. A further degree of error, albeit small,² is introduced as the result of administrative boundary changes that occur between censuses. Although census data can be standardized to take account of boundary changes, annual migration statistics are compiled on the basis of the boundaries as they are each year. The examples used keep such discrepancies to a minimum, and wherever possible corrections have been made.

Volume of Migration

The extent to which census migration data understate the total volume of movement is shown in Table 1. It compares the total number of registered moves between administrative and statistical areas of various levels in two successive decades with the number of moves shown by census data for the same decades.

The discrepancy between the two sources is large. It is clear that for a period of this length, census-derived data obscure more than they reveal, at least in terms of numbers. It is equally clear that this observation applies to internal migration between administrative areas of all levels, where the proportions of hidden movement are remarkably similar, and for both time periods for which data are available, even though migration trends changed around 1970 in Norway as in many other countries. Thus it appears that, when total migration is considered, even though using census data reduces the volume of movement accounted for, the distribution of movement by administrative level, and therefore presumably by distance moved, is not distorted.

Table 1. Levels of Migration According to Data Source and Statistical Level, 1961–1970 and 1971–1980

Statistical unit	Registration (a)	Census (b)	Ratio (a:b)
1961–1970			
Intercommune	1,653,381 ^a	602,451	2.74
Intercounty	890,134	323,117	2.75
Interregion	487,432	173,970	2.76
Between communes, within counties	763,247	279,334	2.73
Between counties, within regions	402,702	149,147	2.70
1971–1980			
Intercommune	1,840,156	675,026	2.73
Inter-trade district	1,376,718	514,714	2.67
Intercounty	969,956	354,610 ^b	2.73
Between communes, within trade districts	463,438	160,312	2.89
Between trade districts, within counties	406,762	160,104	2.54

Note: A list of statistical publications cited with complete references to the sources of data is available from the author. Sources: Census 1980 (1986); County Statistics (1977, 1983); District Statistics (1984); Migration Statistics (1972–1982); Østby (1975, p. 29).

^a See note 1.

^b Total moved to each county. The sum of out-movers from counties was 157 less according to the published tables.

Within-Nation Variation

It does not necessarily follow, however, that the extent to which census data understate migration will be the same in all parts of a given country or in all types of area. The available evidence indicates that they do not.

In Table 2 the level of registered movement is compared with the level of movement shown by census data for each county and for communes within each county grouped by type. Communes are divided into nine categories on the basis of their economic structure, as defined by the sectors in which their inhabitants are employed, population density, and degree of centrality (Stordahl 1983, pp. 65–67). The purpose of the table is to demonstrate, as far as possible with the data available, the extent to which hidden movement varies, not to explain it. That is a separate task.

The comparison is expressed as a ratio of registered to census-derived moves. The higher the ratio, the greater is the proportion of hidden moves. It is evident from the differences between ratios for intercounty and intracounty movement that the similarity of ratios for shorter and longer distance moves in the nation as a whole is not found in all its constituent parts. The variation between counties is greater still, and between communes (for which exchange within the same and with other counties is included), yet greater. The lowest value shown, 2.01 (type 3 in Telemark), indicates that almost half of the moves registered were also evident in the census data. The highest ratio, 3.56 (type 1 in Finnmark), indicates that census data accounted for only slightly more than a quarter (28.1%) of the moves registered over 10 years. This compares with the ratio for all communes, 2.73; that is, overall, 36.5% of registered moves were included in the census data.

The table shows that there is greater variation between counties than between communes of different economic structure and degree of centrality. The highest ratios are found in those counties that experience the highest net out-migration, but there is no equivalent correspondence between counties with the highest net in-migration and low ratios. Thus one cannot conclude that the level of net loss is necessarily related to the level of hidden

Table 2. Ratio of Registered to Census-Derived Moves by County and Commune Type, 1971-1980

County	Commune type									Total		
	1	2	3	4	5	6	7	8	9	Total	Intracounty	Intercounty
Østfold	—	2.17	2.49	2.29	—	2.64	2.79	2.53	—	2.60	2.59	2.60
Akershus	—	2.46	2.36	—	—	2.49	2.74	—	—	2.79	2.89	2.67
Oslo	—	—	—	—	—	—	2.70	—	—	2.70	—	2.70
Hedmark	2.44	2.36	2.47	—	—	2.62	—	2.66	—	2.53	2.61	2.47
Oppland	2.79	2.67	2.49	—	—	2.69	2.39	2.88	2.88	2.69	2.60	2.74
Buskerud	2.57	2.27	2.58	—	—	2.66	2.90	2.54	2.90	2.73	2.73	2.73
Vestfold	—	2.39	2.64	—	—	2.68	—	2.82	—	2.71	2.79	2.62
Telemark	2.66	2.58	2.01	—	2.45	2.71	—	2.70	3.13	2.63	2.63	2.63
Aust-Agder	2.70	2.69	2.54	—	2.69	2.80	2.60	2.92	2.82	2.83	3.08	2.63
Vest-Agder	2.62	2.62	2.84	—	—	2.68	2.83	2.68	—	2.74	2.70	2.77
Rogaland	2.51	2.36	2.91	2.59	2.53	2.34	3.06	2.77	—	2.85	2.85	2.76
Hordaland	2.95	2.52	2.30	2.60	2.97	2.30	3.00	2.76	2.66	2.68	2.55	2.83
Sogn og Fjordane	2.54	2.74	—	2.73	2.34	—	—	3.22	2.99	2.72	2.49	2.74
Møre og Romsdal	2.43	2.69	—	2.75	2.86	—	—	2.86	—	2.82	2.82	2.82
Sør-Trøndelag	2.41	2.58	2.59	2.65	—	—	2.82	—	2.68	2.70	2.52	2.86
Nord-Trøndelag	2.56	2.76	2.62	—	2.60	—	—	2.82	2.75	2.72	2.63	2.77
Nordland	2.58	3.54	—	2.63	2.56	—	—	2.81	2.75	2.72	2.58	2.83
Troms	2.83	2.73	—	2.74	—	—	—	2.88	3.13	2.92	2.63	3.10
Finmark	3.56	—	—	3.53	3.16	—	—	3.11	3.36	3.13	3.08	3.22
Total	2.63	2.63	2.70	2.88	2.68	2.66	2.77	2.80	2.89	2.73	2.56	2.74

Note: The commune types are 1, agricultural communes; 2, noncentral, mixed agricultural, and manufacturing communes; 3, central, mixed agricultural, and manufacturing communes; 4, fishing communes; 5, noncentral manufacturing communes; 6, central manufacturing communes; 7, very central, mixed-service, and manufacturing communes; 8, remaining mixed-service and manufacturing communes; 9, other communes.

Sources: County Statistics (1973-1974, 1983); District Statistics (1984); Migration Statistics (1972-1982).

movement. Rather, it seems more significant that the former counties are also areas where turnover is perceived as a problem, as in the northernmost county, Finnmark (Finstad 1985, pp. 209–219; cf. Morrison 1971, pp. 180–182). Limited though this evidence is, it does seem to suggest that levels of hidden movement might not be related to the same factors as migration gain or loss.

Migration Balance

More important for migration analysis than total movement, in- and out-movement combined, are the differences between the two and thus the difference in the extent to which they are each understated in census-derived data. Though most methods of recording ought to give the same net result, at least where movement in or out of a given area is concerned, the significance of net movement depends on its size relative to the migration streams (Thomas 1936, p. 346). The importance of net movement will be exaggerated in proportion to the amount of movement that is hidden in census-derived data. This follows from the evidence that by far the most important component in hidden movement is repeated movement (Nicholson 1989; cf. Morrison 1971, p. 182); therefore a large proportion of hidden moves cancel each other out.

The extent of the differences can be seen most clearly by comparing ratios of in- to out-movement computed from each source for the same places. The findings shown in Table 3 have been selected from a limited number of examples for which adequate data were available, to illustrate the range of results that have been obtained.

It is evident that not only do census data exaggerate the importance of net movement, but they do so to a degree that varies considerably from one case to another. This is most apparent at the extremes, such as the regions of North and East Norway in the decade of 1961–1970. In general, the difference in the ratio between in- and out-moves, as shown by registration and census-derived data, is least where net migration is so small as to be almost negligible, irrespective of the amount of hidden movement. Thus the differences between the two ratios are very small in the county of Troms, where there is an exceptionally high proportion of hidden movement, and in the southern rural commune of Lardal, where there is relatively little. The county of Finnmark, which has only a slightly greater proportion of hidden movement than Troms, also has a migration deficit, but the high proportion of hidden moves gives an exaggerated impression of this unfavorable position in census data. Even in a rural commune with not excessive hidden movement (Birkenes), the census data suggest that in-migration is almost double out-migration, whereas it is clear from registration data that it is barely 40% higher.

A further problem is demonstrated by the right portion of the table, which shows the ratio of registration to census data for in- and out-movement separately. This is another way of expressing the information shown by the ratios on the left, but one that makes explicit the problems of comparing volumes of in- and out-movement when using census-derived data. The only regularity found is in the ratio of registered to census-derived moves, which is greater in the smaller stream, or, where the two sources differ, the one that according to the census data is the smaller. This is consistent with Eldridge's findings that secondary and return migration, that is, repeat movement, are of greater relative importance in reverse (that is, smaller) than dominant migration streams (Eldridge 1965, pp. 451–452). Similarly, the volumes of in-movement to different places, relative to each other, and out-movement from them, according to the census, will differ from those shown by registration data, given that there is more hidden movement to and from some places than others, as Table 2 showed. Thus the result of comparisons of in- and out-migration may well depend on the source used as much as, perhaps even more than, actual differences in volume.

Table 3. Ratio of In-migration to Out-migration as Shown by Registration and Census-Derived Migration Data and Ratio of Registration to Census Data for In- and Out-Movement

Statistical unit	Ratio, in:out movement		Ratio (a:b)		
	Registration (a)	Census (b)	In	Out	Total
1961-1970					
Regions					
North Norway	.69	.38	4.3	2.4	2.9
East Norway	1.35	2.20	2.3	3.7	2.7
Communes					
Oslo	.95	1.04	2.7	2.9	2.8
1971-1980					
Communes					
Oslo	.83	.75	2.9	2.6	2.7
Sør-Varanger	.85	.66	3.6	2.8	3.1
Vardø	.75	.50	3.9	2.8	3.0
Birkenes	1.37	1.91	2.4	3.1	2.5
Krødsherad	1.52	2.37	2.0	3.1	2.3
Lardal	1.13	1.16	2.3	2.4	2.4
Kristiansand	.94	.90	2.9	2.7	2.8
Counties					
Finnmark	.85	.65	3.7	2.9	3.2
Troms	.99	1.01	3.1	3.1	3.1
Hedmark	1.17	1.27	2.4	2.6	2.5
Telemark	1.05	1.01	2.7	2.6	2.6
Aust-Agder	1.29	1.68	2.4	3.1	2.6

Sources: Census 1980 (1986); County Statistics (1973, 1977, 1983); District Statistics (1984); Migration Statistics (1972-1982); Population Movements (1963-1970, 1978); Statistical Yearbook (1971-1972).

Trends

In a few cases census-derived data misrepresent not only the relative volumes of in- and out-movement but also the direction of net movement. Examples are the city of Oslo between 1960 and 1970 and agricultural communes between 1970 and 1980. Table 4 shows the migration totals and balances given by both sources for each.

In both cases the annual statistics compiled from migration registrations showed that a change in migration trend occurred in the course of the decade. Though the changes

Table 4. Migration Balance as Shown by Registration and Linked Census Data

Data source	In	Out	Net
Oslo, 1961-1971			
Registration	179,186	188,399	-9,213
Linked census data	67,037	64,267	+2,770
Agricultural communes, 1971-1981			
Registration	87,886	85,049	+1,737
Linked census data	32,666	33,200	-534

Sources: County Statistics (1973, 1977, 1983); District Statistics (1984); Migration Statistics (1972-1982); Population Movements (1963-1970); Statistical Yearbook (1971-1972).

became immediately evident in registration data, in the census-derived data the opposing trends within one time period canceled each other out, giving a false impression of continuity and indicating that such changes can remain hidden for several years in census data.

The manner in which change takes place is also subject to distortion in census-derived data, as exemplified by the city of Oslo in the 1960s and the 1970s. Linked census data suggest that there was an abrupt change, whereas registration data for individual years (not shown) show that there was a gradual transition. The data for Oslo in Table 3 also show that the exaggeration of net relative to gross movement in census-derived data results in a corresponding exaggeration of the degree of change in the trend. Thus it is likely that there was a similar exaggeration of the phenomenon of turnaround in counties where migration data are derived from the census (cf. Campbell & Garkovich 1984, pp. 91–92).

Wider Relevance

Though the findings reported here refer to one particular nation and a relatively short time period, research elsewhere demonstrates their general relevance. That hidden moves are unevenly distributed between places and migration streams is a necessary corollary of Morrison's observation in the United States that "variations in metropolitan mobility rates are . . . the product of repeated movement" (Morrison 1971, p. 182). Likewise Eldridge's data suggest that hidden movement is differentially distributed between dominant and reverse migration (Eldridge 1965, pp. 451–452). Rees's comparison of interregional migration in England and Wales, where he compared residence 5 years and 1 year prior to the 1971 census, also indicates variation in the level of hidden movement between streams and regions (Rees 1977, p. 248).

Nor is the scale of hidden movement identified simply due to the length of the intercensal period, which was longer than that usually covered by retrospective census questions on residence. Data for shorter periods (Courgeau 1979, p. 26; Geschwind 1957, p. 614; Nicholson 1989; Rees 1977, p. 248; Tucker & Urton 1987, p. 267) suggest that even over 5 years, up to half of the moves that occur do not appear in census data. These findings also show a remarkable consistency in the level of hidden movement between counties.

Conclusion

Movement that is recorded by a registration system but is hidden in data derived from censuses is clearly no mere "add-on" to all migration streams that can be ignored for most practical purposes. Only when the totality of moves within a nation is considered, as in Table 1, are levels of hidden movement similar, irrespective of the size of statistical unit, and thus, apparently, of distance moved. When data are organized by migration direction (in or out) or by individual streams, as they commonly are, or when total movement in particular areas is considered, levels of hidden movement vary, in some cases considerably.

These variations imply that research findings based on census-derived migration data may be distorted. This applies to all census data on migration, including net values, though it is especially true in the case of individual streams and still more so when considering subgroups within streams. Though it has not been dealt with here, the literature on repeat migration makes it clear that the proportion of hidden movement is likely to vary still more between subgroups of the population, such as age groups (Eldridge 1965, pp. 451–452; Reisz 1976, p. 71; Tucker & Urton 1987, p. 267), than for the population as a whole. It must be a matter of some importance to seek further clarification of the extent of hidden movement in census-derived migration data and the nature of its variation.

Notes

¹ In 1980 there were 454 communes, with from 265 (Utsira) to 451,789 (Oslo) inhabitants and a median population of just under 5,000. There are 18 counties plus the city of Oslo, a county in its own right. The total population of Norway was 4,091,132 (Population 1981, p. 31).

² In the mid-1960s many commune boundaries were changed. Only data that could be adjusted to allow for the changes have been used for this period. In Table 1 the number of moves between communes from 1961 to 1970 was estimated by dividing the number of moves recorded each year by the mobility rate (moves per thousand of the median population), then multiplying by a corrected mobility rate for that year. The corrected mobility rate is the mobility rate that the Central Statistical Bureau estimated would have obtained had commune boundaries in the respective years been as they were on January 1, 1972, when there were 444 communes (Østby 1975, p. 29). The method of calculating the corrected mobility rate was not specified.

Acknowledgments

The volume of data drawn on in this article is due in no small measure to the generosity of the Norwegian Central Statistical Bureau (Statistisk Sentralbyrå). I would like to express my appreciation for their cooperation. Without it this article could not have been written. Willis J. Goudy, Daniel Courgeau, Larry Long, and several anonymous reviewers made valuable comments on earlier drafts.

[Received April 1988. Revised August 1989.]

References

- Byfuglien, J. (1977). "Hva forteller folketellingene 1960 og 1970 om befolkningens skifte av bosted i perioden?" *Plan og Arbeid*, 4, 161-167.
- Campbell, R. R., & Garkovich, L. (1984). "Turnaround Migration as an Episode of Collective Behavior." *Rural Sociology*, 49, 89-105.
- [Census 1980]. (1986). *Folke- og bustadteljing 1980* (Vol. 4). Oslo: Central Bureau of Statistics.
- [County Statistics]. (1973-1974). *Statistisk fylkeshefte 1973* (18 vols.). Oslo: Central Bureau of Statistics.
- . (1977). *Statistisk fylkeshefte 1977* (18 vols.). Oslo: Central Bureau of Statistics.
- . (1983). *Statistisk fylkeshefte 1983* (18 vols.). Oslo: Central Bureau of Statistics.
- Courgeau, D. (1979). "Migrants and Migrations." *Population Selected Papers No. 3*. [Translation of *Population* (1973), 28, 511-537.]
- DaVanzo, J. S., & Morrison, P. A. (1981). "Return and Other Sequences of Migration in the United States." *Demography*, 18, 85-101.
- [District Statistics]. (1984). *Nye distriktstall* (for Akershus and Oslo, Hedmark, and Oppland). Oslo: Central Bureau of Statistics.
- Eldridge, H. T. (1965). "Primary, Secondary, and Return Migration in the United States, 1955-60." *Demography*, 2, 444-455.
- Finstad, N. (ed.). (1985). *Flytting og flyttegrunner—Finnmarksflyttere*, Flytting fra Finnmark Prosjektrapport No. 7. Bodø: Nordlands Distrikthøgskole.
- Geschwind, H. (1957). "Inrikes omflytning 1951-1955." *Statistisk tidskrift*, new ser., 6, 607-614.
- [Migration Statistics]. (1972-1982). *Flyttestatistikk* (annual). Oslo: Central Bureau of Statistics.
- Morrison, P. A. (1971). "Chronic Movers and the Future Redistribution of Population: A Longitudinal Analysis." *Demography*, 8, 171-184.
- Nicholson, B. (1989). "Length of Inter-censal Period and the Scale of Omission in Census Migration Data." Unpublished manuscript, Centre for Scandinavian Studies, Newcastle upon Tyne, UK.
- Østby, L. (1975). *Flyttingene i Norge 1971 og 1949-73*, Statistiske Analyser No. 13. Oslo: Statistisk Sentralbyrå.
- [Population]. (1981). *Folketallet i kommunene 1980-1981*. Oslo: Central Bureau of Statistics.
- [Population Movements]. (1963-1970). *Folkemengdens bevegelse* (annual). Oslo: Central Bureau of Statistics.
- . (1978). *Folkemengdens bevegelse: Oversikt 1971-1975*, Statistiske Analyser No. 36. Oslo: Central Bureau of Statistics.
- Rees, P. H. (1977). "The Measurement of Migration From Census Data and Other Sources." *Environment and Planning*, Ser. A, 9, 247-272.
- Reisz, B. A. (1976). *Flytting og levevilkår*, NIBR Arbeidsrapport 31/76. Oslo: Norsk institutt for by- og regionforskning.

- [*Statistical Yearbook*]. (1971–1972). *Statistisk årbok* (annual). Oslo: Central Bureau of Statistics.
- Stordahl, E. (1983). *Flyttemønstre: Norge 1971–1974*, Samfunnsøkonomiske Studier 50. Oslo: Statistisk Sentralbyrå.
- Thomas, D. S. (1936). "Internal Migrations in Sweden: A Note on Their Extensiveness as Compared With Net Migration Gain and Loss." *American Journal of Sociology*, 42, 345–357.
- Tucker, C. J., & Urton, W. L. (1987). "The Frequency of Geographic Mobility: Findings From the National Health Interview Survey." *Demography*, 24, 265–270.