Sevareid, Eric.
1967 "Dissent or destruction?" Look 31 (September 5):21ff.

Silver, Allan A.
1968 "Official interpretations of racial riots." Robert H. Connery (ed.), Urban Riots: Violence and Social Change. New York: Academy of Political Science.

Simmel, Georg
1955 Conflict and the Web of Group-Affiliations. Translated by Kurt Wolff and Reinhard Bendix. New York: Free Press.

Smelser, Neil J.
1963 Theory of Collective Behavior. New York: Free Press.

Tomlinson, T. M. and David O. Sears.
1967 Los Angeles Riot Study: Negro Attitudes Toward the Riot. Los Angeles: U.C.L.A. Institute of Government and Public Affairs.

Turner, Ralph H.
1957 "The normative coherence of folk concepts." Research Studies of the State College of Washington 25:127–136.
1964 "Collective behavior and conflict: new theoretical frameworks." Sociological Quarterly 5 (Spring):122–132.

Turner, Ralph H. and Lewis M. Killian.
1957 Collective Behavior. Englewood Cliffs, N. J.: Prentice-Hall.

von Wiese, Leopold.
1932 Systematic Sociology. Adapted and amplified by Howard Becker. New York: John Wiley and Sons.

Waller, Willard and Reuben Hill.
1951 The Family: A Dynamic Interpretation. New York: Dryden Press.

ON THE MEASUREMENT OF SOCIAL MOBILITY: AN INDEX OF STATUS PERSISTENCE *

LEO A. GOODMAN
University of Chicago

Considering a cross-classification table that describes an aspect of social mobility (the relation between origin status and destination status) for a population of individuals, this paper shows that the usual indices of mobility (or immobility), which are based upon a comparison of the observed frequencies in the mobility table with the corresponding expected frequencies estimated under the assumption of "perfect mobility," are defective in an important respect. A different index is introduced which is not defective in this respect. For those individuals whose origins are in a given status category, the new index measures the degree to which an individual's status of origin "persists" from his origin to his destination. This index can be used to compare the different status categories of origin with respect to their degree of status persistence, and it can also be used for other comparative purposes. Calculating this index of persistence for the data in the classical studies of intergeneration social mobility in Britain and in Denmark, we find, for example, that (1) its magnitude is negative for those whose origins are in the middle (M) status category (i.e., there is, in a certain sense, an "exodus" from the middle status category); (2) its magnitude is positive for those whose origins are in the upper (U) or lower (L) status categories; (3) it is greater for those whose origins are in the U status category than for those whose origins are in the L status category; and (4) the magnitudes of the index for the different status categories of origin differ from each other in statistically significant ways. The index of persistence introduced here can serve to supplement and extend some of the methods developed in the author's earlier work.

We shall discuss a methodological problem pertaining to the measurement of the degree of social mobility using, for illustrative purposes, five different cross-classification tables describing intergeneration social mobility: (a) two hypothetical 3 x 3 cross-classification tables; (b) two 3 x 3 tables based upon data obtained in the studies of British social mobility by Glass and his co-workers (1954), and of Danish social mobility by Svalastoga (1959); and (c) a 5 x 5 table based upon the data in the British study. The point of view and methods described in the present paper can be applied not only to the analysis of intergeneration social-mobility tables,

* This research was supported in part by Research Contract No. NSF GS 1905 from the Division of the Social Sciences of the National Science Foundation. For helpful comments, the author is indebted to J. Fennessey, S. Fienberg, S. Haberman, D. McFarland, and T. Pullum.

but also to the analysis of certain kinds of intrageneration social-mobility tables, and more generally to the analysis of cross-classification tables that describe certain kinds of processes of change.

Much has been written on methodological problems pertaining to the measurement of social mobility (see, e.g., Yasudo, 1964; Duncan, 1966; Wilensky, 1966). In the present paper, we shall begin with a cross-classification table that describes an aspect of social mobility (the relation between origin status and destination status) for a population of individuals. We shall first focus our attention on a problem that arises in the interpretation of any of the usual indices of this aspect of social mobility (or immobility) calculated from the data in the cross-classification table. We shall show that these indices (which are based, in one way or another, on a comparison of the observed frequencies in the table with the corresponding "expected frequencies" estimated in the usual way) [1] can lead to incorrect or misleading interpretations of the data. To remedy this defect, we shall suggest that the observed frequencies in the table should be compared with a different kind of "expected frequency"; [2] and after making this comparison, we shall consider the problem of measuring the degree to which an individual's status of origin "persists" from his origin to his destination. We shall introduce herein an index that measures, for any given origin status, the degree to which the status of origin "persists" for those people having the given origin status; and we shall compare this index with some of the other measures that have been proposed in the earlier literature on social mobility. We shall also show how to apply some of the statistical methods developed in an earlier article on the analysis of "persistence," by the present author (Goodman, 1964b), to test whether the degree of status persistence is different for people who differ with respect to their origin status.

TWO SIMPLE EXAMPLES

Let us consider the hypothetical cross-classification described by Table 1A. Because the data in this social-mobility table are of particularly simple form, we can analyze these data by sight alone, ignoring (for the moment) even the methods of analysis taught in introductory courses in statistics.[3] By examining Table 1A by sight alone, we note that (1) people with origin status L are distributed uniformly in the three status categories of destination; (2) people with origin status M are also distributed uniformly except for a deficiency

TABLE 1. CROSS-CLASSIFICATION OF TWO HYPOTHETICAL SAMPLES OF MALES ACCORDING TO EACH SUBJECT'S STATUS CATEGORY (CATEGORY OF DESTINATION) AND HIS FATHER'S STATUS CATEGORY (CATEGORY OF ORIGIN)

Table 1A: First Sample

		Subject's Status		
		U	M	L
Father's Status	U	144	36	36
	M	360	252	360
	L	36	36	36

Table 1B: Second Sample

		Subject's Status		
		U	M	L
Father's Status	U	250	300	200
	M	200	600	400
	L	100	300	200

[1] The usual "expected frequencies" are estimated under the assumption that there is "perfect mobility" from origin status to destination status. The assumption of "perfect mobility" means that each individual's destination status (i.e., his column classification) is assumed to be independent (in a statistical sense) of his origin status (i.e., his row clasification), for the population of individuals to which the cross-classification table pertains.

[2] The kind of "expected frequency" that we recommend will be estimated under a more realistic assumption than the usual "perfect mobility" assumption. For this purpose, the more general concept of "quasi-perfect mobility" (or "quasi-independence") was introduced, and appropriate methods for estimating the "expected frequencies" under the assumption of "quasi-perfect mobility" were developed, in an earlier series of articles by the present author (Goodman, 1963, 1964a, 1965, 1968, 1969).

[3] Indeed, if the reader will erase from his mind (for the time being) the methods of analysis he has been taught in introductory courses in statistics, he will probably find it easier to follow us in our analysis by sight of the data in Table 1A.

of individuals with destination status M; and (3) people with origin status U are also distributed uniformly except for an excess of individuals with destination status U. This view of Table 1A would suggest that there is (1) no tendency for origin status L to persist, (2) a tendency for there to be an "exodus" from origin status M (i.e., for persistence to be "negative" from origin status M), and (3) a tendency for origin status U to "persist" (i.e., for persistence to be "positive" for origin status U).

We have introduced Table 1A herein because it has the simple interpretation noted above,[4] and because it can be used as an indicator for judging which of the various indices of interest are compatible, and which are incompatible, with this interpretation.[5] If the numerical value of a given index calculated for Table 1A is incompatible with our interpretation of this simple table, we would then be doubtful about the utility of the index for the analysis of less simple mobility tables (e.g., the social-mobility tables for Britain and Denmark described in the next section).[6] We shall see later herein that Table 1A is similar, in some important respects, to the 3 x 3 tables describing social mobility in Britain and in Denmark.[7] This fact adds some weight to the use of Table 1A as an indicator for judging the utility of the various indices of interest.[8]

Let us now reexamine Table 1A using the index of social immobility based upon the usual mobility ratios (i.e., the ratios of the observed frequencies in the table and the corresponding expected frequencies estimated under the assumption of perfect mobility).[9] If there were perfect mobility, then the usual estimate of the expected frequencies in the three diagonal cells of Table 1A would be 90, 243, and 36, for the (U,U), (M,M), and (L,L) cells, respectively. Comparing the observed frequencies in these cells (viz., 144, 252, and 36) with the corresponding expected frequencies, as is usually done, we obtain ratios of 144/90=1.60, 252/243 =1.04, and 36/36=1.00, respectively. Thus, by the usual methods, we see that the observed frequency in the diagonal cells of Table 1A is greater than the corresponding expected frequency (estimated under the assumption of perfect mobility) for origin status categories U and M; and the observed and expected frequencies are equal for origin status L. The reader might get the impression from the interpretation of the mobility ratios that (a) there is a tendency for status origins U and M to persist from origin to destination, and (b) there is no such tendency for origin status L. Thus, the use of the usual mobility-ratio methods would convey an impression concerning mobility from origin status M that is incompatible with the simple interpretation arrived at by sight earlier in this section.

Let us now examine why the interpretations were incompatible. The interpretation in the preceding paragraph uses the "perfect-mobility" model to obtain a standard with which to compare the observed data in

---

[4] Of course, other interpretations of this table are possible. However, for Table 1A, the simple interpretation obtained by sight alone is entirely congruent with the data, and it leads to an explanatory model for the table that is, in a certain sense, complete. We shall discuss this further later herein.

[5] For a given mobility table, two indices are described as "incompatible," if the general impression pertaining to some aspect of mobility, which is conveyed by the interpretation of one of the indices, appears to contradict the corresponding general impression conveyed by the interpretation of the other index. Since different indices actually measure different things, a comparison of the indices will help to clarify what each is in fact measuring.

[6] On the other hand, it should be noted that an index that is defective in the sense described above may, nevertheless, have other merits.

[7] I noted in my earlier article (1965) that there was "quasi-perfect mobility" in the British and Danish tables, when the entries in the three diagonal cells are blanked out (i.e., when status inheritance in the three status categories is taken into account). We shall see below that there is also "quasi-perfect mobility" in Table 1A, when the entries in the three diagonal cells are blanked out. (There is also "quasi-perfect mobility" in Table 1A, when the entries in the first two diagonal cells—(U,U) and (M,M)—are blanked out.)

[8] Other hypothetical tables that (a) have simple interpretations and (b) are similar, in some important respects, to actual social-mobility tables, could also be used as indicators for judging the utility of the various indices. Different hypothetical tables might lead to different judgments about a particular index.

[9] See footnote 1.

Table 1A; whereas, the interpretation arrived at by sight uses a different kind of model [10] to obtain a different standard for purposes of comparison. The observed data contradict the "perfect mobility" assumption; [11] but they do not contradict the model pertaining to the interpretation arrived at by sight.[12] In arriving at a standard with which to compare the data, use of the model that does not contradict the data will usually be preferable to use of the unrealistic "perfect-mobility" model.

To gain further insight into the problem of interpretation that arises when the "perfect-mobility" model is used to obtain the standard of comparison, let us consider for the moment a different hypothetical mobility table; viz., Table 1B. Note that if the entry in the (U,U) cell of this table had been 100, rather than 250, there would have been "perfect mobility" in the table. Since Table 1B could serve as an example of "perfect-mobility" simply by reducing the entry in only one of its cells (viz., the (U,U) cell), we find it worthwhile to use this table, as well as Table 1A, as an indicator for judging the utility of the usual indices based upon the "perfect-mobility" model.

First, let us examine Table 1B by sight alone. Having noted that, aside from the excess of people in the (U,U) cell, there is "perfect mobility" in this table, we would conclude that there is (1) no tendency for status origins M or L to persist, and (2) a tendency for status origin U to persist. On the other hand, a reexamination of this table using the usual index of status immobility would convey an impression concerning mobility from status origins M and L that is incompatible with the simple interpretation noted above. We shall now explain why this would happen.

The defect noted above in the interpretation of Table 1B obtained with the usual immobility index (based upon the mobility ratios) is due to the fact that, in calculating the usual expected frequencies, if the observed frequency in cell (U,U) is "too large" (in the sense described above), this large frequency will increase the column and row marginal totals pertaining to category U, which will in turn decrease the relative size of the column and row marginals pertaining to the remaining categories (i.e., categories M and L), relative to category U. With this relative decrease in the marginals pertaining to categories M and L, we obtain a decrease in the expected frequencies in cells (M,M) and (L,L) (estimated under the assumption of perfect mobility); and thus, the ratios of the observed to the expected frequencies in the (M,M) and (L,L) cells are raised above 1. But the fact that these mobility ratios are greater than 1 would convey an impression [13] about Table 1B that would contradict the conclusion arrived at earlier, by sight alone, that there is no tendency for status origins M or L to persist.

Many of the usual measures of social mobility are based, in one form or another, on the comparison of the observed frequencies with the corresponding expected frequencies estimated under the assumption of perfect mobility. (See, e.g., Rogoff, 1953; Glass, 1954; Carlsson, 1958; Svalastoga, 1959.) Considering those measures that are based upon this kind of comparison, we have noted that they are subject to the following defect: The estimate of the expected frequency in any given diagonal cell, say the (M,M) cell, is affected by an observed excess (or, possibly, a dearth) of individuals in any other diagonal cell, say the (L,L) cell.[14]

---

[10] We shall discuss this model more fully in the next section.

[11] For Table 1A, an individual's destination status is obviously dependent upon his origin status.

[12] See further discussion in the next section.

[13] This kind of defective impression, which is obtained with the mobility ratios, would be obtained whenever the "perfect mobility" model is applied as a standard to any cross-classification table for which, aside from the excess of people in the (U,U) cell, there is "perfect mobility" in the table.

[14] The two hypothetical tables (Tables 1A and 1B), considered in this section, and the British and Danish 3 x 3 mobility tables (Tables 2A and 2B), considered in the following sections, have an observed excess (or a dearth) of individuals only in some (or all) of the diagonal cells; and, as we shall see in the next section, there is "quasi-perfect mobility" in the non-diagonal cells. Because of this, we have not been concerned, at this point in the exposition, with the case where an observed

This defect can lead to an incorrect or misleading interpretation of the data.

QUASI-PERFECT MOBILITY

We noted earlier that, if the entry in the (U,U) cell of Table 1B were changed from 250 to 100, there would be "perfect mobility" in the modified table. Because of this, we say that there is "quasi-perfect mobility" in Table 1B when the entry in the (U,U) cell is "adjusted." Letting R_j denote the probability that an individual will fall in destination status category j (for j=1, 2, 3 corresponding to U, M, L, respectively), we see that $R_1=1/6$, $R_2=1/2$, and $R_3=1/3$ for the modified form of Table 1B. Letting P_i denote the probability that an individual will fall in origin status category i (for i=1, 2, 3 corresponding to U, M, L, respectively), we see that $P_1=1/4$, $P_2=1/2$, and $P_3=1/4$ for the modified form of this table. Note that

$$\sum_{j=1}^{3} R_j=1 \text{ and } \sum_{i=1}^{3} P_i=1, \qquad (1)$$

for the 3 x 3 table. Letting P_{ij} denote the probability that an individual will fall in origin status i and destination status j, we see that

$$P_{ij}=P_iR_j, \qquad (2)$$

for each cell (i,j) in the modified table. Equation (2) states that, if we consider the individuals in the modified table, an individual's destination status is independent (in a statistical sense) of his origin status; i.e., that there is "perfect mobility" in the modified table.

Let us now consider Table 1A. If the entries in cells (U,U) and (M,M) of this table are changed from 144 and 252 to 36 and 360, respectively, there would be "perfect mobility" in the modified table. Because of this, we say that there is "quasi-perfect mobility" in Table 1A when the entries in cells (U,U) and (M,M) are "adjusted." For the modified form of this table, we find that $R_1=R_2=R_3=1/3$, and $P_1=1/12$, $P_2=5/6$, $P_3=1/12$; and equation (2) is satisfied when these values of R_j and P_i are used.

We have defined "quasi-perfect mobility" here for a given table (say, Table 1A or 1B) by the condition that equation (2) is satisfied for a modified form of the table in which the entries in certain specified cells have been adjusted. An alternative, but equivalent, definition of "quasi-perfect mobility" is the following: Let us blank out the entries in certain specified cells of the table (e.g., cell (U,U) in Table 1B, and cells (U,U) and (M,M) in Table 1A). Considering now the individuals in the table with its blanked-out cells, let P°_{ij} denote the probability that an individual will fall in origin status i and destination status j (i.e., in cell (i,j)) of this table. We assign a zero probability to each blanked-out cell. Then there is "quasi-perfect mobility" in the table if the P°_{ij} satisfy the following equation for each cell (i,j) that is not blanked out:

$$P^\circ_{ij}=P_iR_j/[\Sigma^*P_iR_j], \qquad (3)$$

where the symbol $\Sigma^*P_iR_j$ denotes the summation of P_iR_j over all cells (i,j) that are not blanked out. Note that equation (3) is satisfied for Table 1B when cell (U,U) is blanked out; and also for Table 1A when cells (U,U) and (M,M) are blanked out.[15] In equation (3), the parameters R_j and P_i are positive constants that will satisfy this equation, and that also satisfy equation (1).

Let us now consider the case where all the diagonal cells in the table are blanked out. In this case, equation (3) can be rewritten as

$$P^\circ_{ij}=P_iR_j/[1-\sum_{k=1}^{3} P_kR_k], \qquad (4)$$

for all cells (i,j) where $i\neq j$. Here we have excluded from consideration (i.e., we have blanked out) any individual whose origin

excess (or a dearth) of individuals occurs also in some non-diagonal cells. (This topic will be discussed in the later section on the analysis of the British 5 x 5 mobility table (Table 7).) As we noted earlier, we have been concerned here with particularly simple examples (in which there is "quasi-perfect mobility" in the non-diagonal cells) because, if an index leads to a defective interpretation for these simple tables, we would be doubtful about the utility of the index calculated for less simple mobility tables.

[15] Furthermore, equation (3) is satisfied for Table 1B when any other cells (for example, (M,M) and (L,L)) are blanked out in addition to cell (U,U); and also for Table 1A when any other cells (for example, (L,L)) are blanked out in addition to cells (U,U) and (M,M).

and destination status categories are the same.

To gain a more complete understanding of the meaning of the R_j and P_i, we first return to the definition of "quasi-perfect" mobility using the "adjusted" entries (rather than "blanked-out" entries), which we introduced at the beginning of this section. If we now focus our attention only on the non-diagonal entries of the table, and allow the diagonal entries to be "adjusted" in a way that leads to "perfect mobility" in the modified table, then the probability P_{ij} (defined earlier herein) for the modified table will satisfy equation (2) where R_j denotes the probability that an individual (in the modified table) will fall in destination status category j, and P_i denotes the probability that an individual (in the modified table) will fall in origin status category i.[16] If we consider now the table in which the diagonal cells have been blanked out, then the probability P°_{ij} (defined earlier herein) for this table will satisfy (4), where the R_j and P_i are the quantities defined above for the modified table.[17] Thus, in this case, we might describe R_j as the "theoretical tendency" for an individual to fall in the destination status category j, considering only those individuals whose origin and destination status categories are different; and a similar kind of description applies to the P_i. In other words, having excluded from consideration the individuals in the table whose origin and destination status categories are the same, R_j is the hypothetical proportion of individuals in destination status category j in the population described by the corresponding modified table (i.e., in the hypothetical population in which the diagonal cells have been "adjusted" to obtain "perfect mobility"[18]). A similar kind of definition can be given for the P_i.[19] (A somewhat different, and perhaps simpler, interpretation for the R_j will be presented in the following section.)

At the beginning of the present section, we calculated the numerical values of R_j and P_i for Tables 1A and 1B, and we noted that there is "quasi-perfect mobility" in these tables. These calculations were elementary because of the particularly simple form of these tables. We shall now consider tables that have a somewhat less simple form.

TABLE 2. CROSS-CLASSIFICATION OF BRITISH AND DANISH MALE SAMPLES ACCORDING TO EACH SUBJECT'S STATUS CATEGORY (CATEGORY OF DESTINATION) AND HIS FATHER'S STATUS CATEGORY (CATEGORY OF ORIGIN)

Table 2A: British Sample

		Subject's Status		
		U	M	L
	U	588	395	159
Father's Status	M	349	714	447
	L	114	320	411

Table 2B: Danish Sample

		Subject's Status		
		U	M	L
	U	685	280	118
Father's Status	M	232	348	198
	L	83	201	246

Let us turn to the two mobility tables in Table 2. These tables provide cross-classifications of male samples in Britain and in Denmark according to each subject's occupational status category and his father's occupational status category. The U, M,

[16] If we focus our attention only on the non-diagonal entries of Table 1A, and allow the diagonal entries to be "adjusted" in a way that leads to "perfect mobility" in the modified table, then the entries in the diagonal cells (U,U), (M,M), (L,L) in the modified table will be 36, 360, 36 respectively; i.e., the adjustment actually changes the entries in only two out of the three diagonal cells (viz, (U,U) and (M,M)). Similarly, if we focus our attention only on the non-diagonal entries of Table 1B, the adjustment of the diagonal entries actually changes only one out of the three diagonal entries.

[17] We use the term "modified table" to refer to the table in which certain entries have been "adjusted" in the way indicated above. We do not use this term to refer to the table in which certain cells have been blanked out.

[18] Since there is "perfect mobility" in the modified table, the same hypothetical proportion R_j (for $j=1,2,3$) applies for individuals in each of the origin status categories in the modified table. In other words, R_j is equal to the hypothetical proportion of individuals in destination status category j in the population described by the individuals who are in origin status category i in the modified table (for $i=1,2,3$).

[19] See my article (1968) for further comments.

and L status categories for the British data correspond to the occupational status categories 1–4, 5, and 6–7, respectively, as defined by Glass and his co-workers (1954); and for the Danish data they correspond to occupational categories 1–6, 7, 8–9, respectively, as defined by Svalastoga (1959).[20]

In my earlier article (1965), I showed that the observed pattern of frequencies in the non-diagonal cells of these cross-classification tables was congruent with the thesis that there was quasi-perfect mobility in the non-diagonal cells. This was done for each of these tables by (a) blanking out the diagonal cells in the table; (b) calculating estimates $\hat{R}_j$ and $\hat{P}_i$ of the parameters R_j and P_i under the assumption of quasi-perfect mobility in the non-diagonal cells;[21] (c) calculating estimates $\hat{F}^\circ_{ij}$ of the "expected frequencies" in the non-diagonal cells under this assumption of quasi-perfect mobility, using the formula[22]

$$\hat{F}^\circ_{ij} = n^\circ \hat{P}^\circ_{ij}, \qquad (5)$$

where n° is the number of individuals in the non-diagonal cells and $\hat{P}^\circ_{ij}$ is the estimate of P°_{ij} calculated by replacing R_j and P_i in equation (4) by their corresponding estimates $\hat{R}_j$ and $\hat{P}_i$; and then (d) comparing the observed frequency f_{ij} in the non-diagonal cells ($i \neq j$) with the corresponding $\hat{F}^\circ_{ij}$ using the chi-square goodness-of-fit statistic[23]

TABLE 3. EXPECTED MOBILITY PATTERN FOR BRITISH DANISH MALE SAMPLES WHEN QUASI-PERFECT MOBILITY IS ASSUMED FOR THOSE INDIVIDUALS WITH DESTINATION STATUS CATEGORIES THAT DIFFER FROM THEIR ORIGIN STATUS CATEGORIES

Table 3A: British Sample

		Subject's Status		
		U	M	L
Father's Status	U	588*	390.2	163.8
	M	353.8	714*	442.2
	L	109.2	324.8	411*

Table 3B: Danish Sample

		Subject's Status		
		U	M	L
Father's Status	U	685*	284.7	113.3
	M	227.3	348*	202.7
	L	87.7	196.3	246*

* These cells were blanked out when calculating the expected mobility pattern for the cells that were not blanked out; i.e., for the non-diagonal cells. The frequencies, which are given here in the cells that were blanked out, are the observed frequencies.

$$X^2 = \Sigma^* (f_{ij} - \hat{F}^\circ_{ij})^2 / \hat{F}^\circ_{ij}, \qquad (6)$$

where the summation is over all non-diagonal cells.[24]

The numerical values of the $\hat{F}^\circ_{ij}$, obtained by the method described above and (in more detail) in the Appendix A1 herein, are given in Table 3.[25] Applying (6) to compare the f_{ij} in Table 2 with the $\hat{F}^\circ_{ij}$ in Table 3, we obtain $X^2 = 0.6$ for the British

[20] These tables were studied earlier by White (1963) and Goodman (1965).

[21] The calculation of the $\hat{R}_j$ and $\hat{P}_i$ is described (in summary form) in Appendix A1 herein. These methods were developed for the case where the data in the mobility table describe either (1) a simple random sample of individuals or (2) a stratified random sample in which the column categories (or the row categories) of the table form the strata that are sampled. We use these methods (and related methods) here as an approximate gauge even though the data in Table 2 were actually obtained by a kind of stratified sampling different from that described above.

[22] As was noted in my earlier work (Goodman, 1963, 1964a), we actually can calculate $\hat{F}^\circ_{ij}$ more directly without first calculating $\hat{R}_j$, $\hat{P}_i$, and $\hat{P}^\circ_{ij}$; see also Appendix A1 herein.

[23] In addition to the chi-square goodness-of-fit statistic, my article (1965) gave several other methods for comparing the observed frequencies with the "expected frequencies." For the sake of simplicity and brevity, we shall not discuss these other methods here, though they are useful.

[24] In other words, X^2 is obtained by summing $(f_{ij} - \hat{F}^\circ_{ij})^2/\hat{F}^\circ_{ij}$ over all cell (i,j) that are not blanked out, i.e., over all non-diagonal cells.

[25] There should, of course, be blanks in the diagonal cells of Table 3; but we have inserted the observed frequencies in the cells that were actually blanked out. In the cells that were actually blanked out in Table 17 of my earlier article (1969), the "adjusted" frequencies (rather than observed frequencies) were inserted; i.e., the "frequencies" that would make the expected mobility pattern conform to a pattern of "perfect mobility" for the entire cross-classification table. Thus, Table 3 above (with blanks in the diagonal cells) gives the estimated expected frequencies under quasi-

sample, and $X^2=0.8$ for the Danish sample. Since each of the chi-square statistics has one degree of freedom under the null hypothesis of quasi-perfect mobility, we are impressed with how well the observed frequencies in these samples can be fitted under the assumption of quasi-perfect mobility.[26]

Having noted that the assumption of "quasi-perfect mobility" is more realistic than the more usual assumption of "perfect mobility," we shall now introduce an index that uses as a standard this more realistic assumption. We shall first consider the case where all diagonal cells are blanked out (as above), and then the case where only some of the diagonal cells are blanked out.[27] In a later section, we shall consider the case where some non-diagonal cells are blanked out in addition to the diagonal cells.

AN INDEX OF PERSISTENCE

In Appendix A1, we give a method for calculating the estimates $\hat{R}_j$ and $\hat{P}_i$ of the corresponding parameters R_j and P_i, under the assumption of quasi-perfect mobility.[28] Applying these methods to the British and Danish samples, we obtain the following numerical values for $\hat{R}_1$, $\hat{R}_2$, and $\hat{R}_3$: 0.19, 0.57, and 0.24, respectively, for the British data; and 0.24, 0.54, and 0.21, respectively, for the Danish data. (All calculations presented in this paper were carried out to more significant digits than are reported here.) Since $\hat{R}_j$ estimates the theoretical tendency for an individual to fall in destination status category j (calculated from the entries in the non-diagonal cells), it would seem natural to compare $\hat{R}_j$ with the observed proportion $\hat{A}_j$ of individuals who fall in destination status category j among those whose origin status category was j. Since the maximum possible value of $\hat{A}_j$ is 1 (in the case of complete persistence), we shall compare $\hat{A}_j-\hat{R}_j$ with $1-R_j$, thus obtaining the following index of persistence:[29]

$$\hat{G}_j=(\hat{A}_j-\hat{R}_j)/(1-\hat{R}_j), \text{ for } j=1,2,3. \quad (7)$$

Table 4 gives the numerical values of the index of persistence $\hat{G}_j$ for the British and Danish samples. Note that (1) $\hat{G}_j$ is negative for the status category M, and it is positive for the status categories U and L, in both the British and Danish samples; (2) $\hat{G}_j$ is larger for status category U than for status category L, in both the British and Danish samples; (3) $\hat{G}_j$ is larger for status category U in the Danish sample than in the British sample, and for the other two status categories the corresponding values in the two samples are approximately equal.

TABLE 4. AN INDEX OF THE DEGREE TO WHICH AN INDIVIDUAL'S STATUS OF ORIGIN PERSISTS FROM HIS ORIGIN TO HIS DESTINATION, CALCULATED FOR EACH STATUS CATEGORY OF ORIGIN IN THE BRITISH AND DANISH SAMPLES

Status Category	British Sample	Danish Sample
U	0.40	0.52
M	−0.22	−0.21
L	0.32	0.32

perfect mobility corresponding to the observed frequencies (with the diagonal cells blanked out) in Table 2 herein; and Table 17 of my earlier article (1969) gives the "modified table" (in the sense described in footnote 17 herein) corresponding to Table 3 above. (A typographical error in Table 17: "358.8" should be "353.8".)

[26] See footnote 23.

[27] Recall the earlier discussion of Tables 1B and 1A.

[28] The parameters R_j and P_i were defined in the preceding section.

[29] Other names for this index of "persistence" may in some respects be preferable and in other respects not. In any case, the meaning of the index is clear. Since our data give only origin and destination status categories, we are unable (with these data) to measure or make inferences about status changes that might have taken place during the period from the individual's "origin" to the time his "destination" was established. Of course, the fact that an individual is in the same origin and destination status categories does not imply that he remained continuously in the status category from his "origin" to the time his "destination" was established; he may have moved several times. Our index of "persistence" reflects an aspect ot the observed relationship between origin and destination status categories.

The index $\hat{G}_j$ measures the difference between the observed proportion $\hat{A}_j$ and $\hat{R}_j$, relative to the difference between 1 and $\hat{R}_j$. It is a normed index in the sense that its maximum possible value is one. If the $\hat{R}_j$ are less than a constant C, then the minimum possible value of the index is $-C/(1-C)$. We shall now provide some further insight into the interpretation of $\hat{G}_j$.

Let π_{ij} denote the probability that an individual's destination status will be j, given that his origin status is i. From the definition of π_{ij} for the 3 x 3 table, we see that

$$\sum_{j=1}^{3} \pi_{ij}=1, \text{ for } i=1,2,3. \qquad (8)$$

In my earlier article (1965), I noted that the thesis of quasi-perfect mobility for the non-diagonal cells could be described by equation (4) presented earlier herein, or equivalently, by the following equation:

$$\pi_{ij}=\begin{cases} A_i, & \text{for } i=j \\ (1-A_i)R_j/(1-R_i), & \text{for } i\neq j, \end{cases} \qquad (9)$$

where R_j is the theoretical tendency described in the preceding section, and where A_i is the probability that an individual's destination status will be i given that his origin status is i. Writing

$$G_j=(A_j-R_j)/(1-R_j), \qquad (10)$$

we find that equation (9) can also be written as

$$\pi_{ij}=\begin{cases} G_i+(1-G_i)R_i, & \text{for } i=j \\ (1-G_i)R_j, & \text{for } i\neq j. \end{cases} \qquad (11)$$

Note that our index $\hat{G}_j$ given by equation (7) is an estimate of G_j given by equation (10).

Equation (11) can be interpreted as follows: Let us suppose that the population of individuals who have origin status i can be divided into two groups: a "stayer" group and a "mover" group. For an individual in the "stayer" group, the probability is 1 that his destination status will be the same as his origin status; [30] and for an individual in the "mover" group, the probability is R_j that his destination status will be j (for j=1,2,3).[31] For the population of individuals who have origin status i, let G_i be the proportion of them who are in the "stayer" group, and let $1-G_i$ be the proportion of them who are in the "mover" group. Considering now each individual from this population (i.e., from the population of those who have origin status i), we find that the probability π_{ij} that his destination status will be j can be described by formula (11). Thus, since formula (11) was first obtained above by rewriting equation (9) using the definition of G_j given by formula (10), we see that the quantity G_j defined by formula (10) can be interpreted simply as the proportion of "stayers" among those having origin status j.[32] Similarly, the parameter R_j, which we first defined in the preceding section, can be interpreted simply as the probability (for a "mover") that his destination status will be j (for j=1,2,3).[33]

The following remarks will provide an interpretation of G_j for the case where G_j is negative; i.e., when $A_j<R_j$. In this case, let H_j denote the negative of G_j.[34] Writing

$$H_j=-G_j, \qquad (12)$$

we find that equation (11) can be rewritten as

$$\pi_{ij}=\begin{cases} R_i-H_i\ (1-R_i), & \text{for } i=j \\ (1+H_i)R_j, & \text{for } i\neq j. \end{cases} \qquad (13)$$

---

[30] If the destination status for a given individual happens to be the same as his origin status, this does not necessarily mean that he is in the "stayer" group (unless the probability of this having happened to him was one). An individual whose probability (of having the same destination status as origin status) was less than one may nevertheless find that his destination status happens to be the same as his origin status.

[31] For a "mover," the probability R_j that his destination status will be j does not depend upon his origin status. For further discussion of the mover-stayer model, see Blumen, Kogan, and McCarthy (1955), and Goodman (1961).

[32] Having provided a simple interpretation for G_j, the quantity $\hat{G}_j$ defined by formula (7) can be interpreted in a similar way; viz., as the estimated proportion of "stayers" among those having origin status j. This provides some further clarification of the term "persistence" as used herein.

[33] Compare this simple interpretation of R_j with the more complicated interpretation of this parameter in the preceding section. The estimate $\hat{R}_j$ can also be interpreted in a similarly simple way.

[34] Of course, when G_j is negative, H_j will be positive.

Equation (13) can be interpreted as follows: For the population of individuals who have origin status i, let us suppose that their destination status is determined in two stages. At the first stage, for an individual in the population who had origin status i, the probability is R_j that his status category will be j (for $j=1,2,3$) at this stage.[35] After the first stage, a proportion H_i of the individuals whose origin status was i are "removed" from among those whose status category was i at the first stage;[36] and, for each of the individuals who are "removed," the probability is R_j that his destination status at the second stage will be j (for $j=1,2,3$).[37] For each individual who was not "removed" after the first stage, his status category at the first stage becomes his destination status at the second stage.[38] Considering now each individual who has origin status i, we find that formula (13) describes the probability π_{ij} that his destination status at the second stage will be j. Thus, since formula (13) was first obtained above by rewriting equation (11) using the definition of H_j given by formula (12), we see that the quantity H_j defined by formula (12) can be interpreted simply as the proportion of those in origin status j who will be "removed" from this status category after the first stage. This serves as an interpretation of $-G_j$ when G_j is negative.[39]

The two stages (and the concepts related to them) in the two-stage model described in the preceding paragraph can be viewed as generalized abstractions; they need not be viewed as specific and concrete phenomena. The two-stage model refers to the various factors (psychological, sociological, genetic, *etc.*) that may lead to a decrease in the chances that an individual's status category would be the same as his father's.

We have described in this section two different statistical models: (a) the mover-stayer model, and (b) the two-stage model. When there is quasi-perfect mobility in the non-diagonal cells, model (a) will fit the data for each origin status category j for which $\hat{G}_j$ is positive, and model (b) will fit the data for each origin status category j for which $\hat{G}_j$ is negative. For both the British and Danish samples (Table 2), model (a) is relevant for origin status categories U and L, and model (b) is revelant for origin status category M.

Before closing this section, we note that, although the methods in this section were described for the case where all the diag-diagonal cells are blanked out (and there is quasi-perfect mobility in the non-diagonal cells), they can be easily extended to the case where only some of the diagonal cells are blanked out (and there is quasi-perfect mobility in the cells that are not blanked out).[40] In this case, we see that (a) the $\hat{R}_j$ would be calculated using the data in the cells that are not blanked out; and (b) for any origin status for which the diagonal cell was not blanked out, the numerical value of the corresponding $\hat{G}_j$ would be zero (or close to zero) if there was quasi-perfect mobility in the cells that were not blanked out.

---

[35] The probability R_j that his status category will be j at the first stage does not depend upon his origin status.

[36] If we consider, say, origin status $i=1$, the proportion of individuals in the population in this origin status who will be in the same status category *at* the first stage is R_1, and the proportion who will be in a different status category *at* the first stage is $R_2+R_3=1-R_1$. The individuals referred to above who are "removed" from status category $i=1$ *after* the first stage are selected from among those whose origin status was $i=1$ and whose status category *at* the first stage was $i=1$.

[37] If we consider, say, origin status $i=1$, for an individual who was "removed," the probability is R_1 that his destination status at the second stage will be $i=1$; and the probability is $R_2+R_3=1-R_1$ that his destination status will be $j\neq 1$.

[38] Considering those individuals whose origin status was i, the individuals who were not "removed" include (1) those whose status category at the first stage differed from their origin status, and (2) those whose status category at the first stage was the same as their origin status but who were not among those who were "removed."

[39] A comment similar to footnote 32 can be applied here.

[40] Recall Tables 1A and 1B. We are assuming here that none of the non-diagonal cells have been blanked out. The case where some of the non-diagonal cells are also blanked will be discussed in a later section herein.

COMPARISON OF DIFFERENT KINDS OF INDICES

Table 5 compares the index introduced in the preceding section with two other kinds of indices calculated for the British and Danish data in Table 2. First, let us discuss the comparison of our persistence index with the index of immobility based upon the usual mobility ratio. We found earlier that persistence was negative for origin status M in both the British and Danish samples; i.e., there was a tendency for there to be an "exodus" from origin status M. Now we find that the usual immobility index for this status category is larger than one. Thus, we note that (as in our earlier discussion of Tables 1A and 1B) the impression conveyed by the usual immobility index is incompatible with our interpretation of the persistence index. The usual immobility index was unable to detect the negative persistence, nor was it able to detect a case of zero persistence.[41]

Let us now consider the index of "conditional uncertainty" suggested by McFarland (1969) for the analysis of mobility tables.[42] For the population of individuals who were in origin status category i, the "conditional uncertainty" index H_i describes an aspect of the distribution (π_{i1}, π_{i2}, π_{i3}) of these individuals in the various destination status categories (j=1,2,3). This aspect indicates how different (in a certain special sense) this distribution is from two particular kinds of extreme distributions: (1) the uniform distribution (in which $\pi_{i1}=\pi_{i2}=\pi_{i3}=1/3$) describing "maximum uncertainty," and (2) the distributions describing "minimum uncertainty" in which one of three π's is 1 and the other π's are zero (e.g., $\pi_{i1}=1$ and $\pi_{i2}=\pi_{i3}=0$). It should be noted that (a) this index does not distinguish in any way between the destination status category that is the same as the origin status category (i.e., the status category j, with j=i) and the destination status categories that are different from the origin status categories (i.e., the status categories j, with $j\neq i$);[43] (b) the index of "uncertainty" can give the same numerical value when calculated for two distributions that differ from each other in important respects (e.g., one distribution might provide evidence of positive status persistence, and the other distribution might provide evidence of negative status persistence, in an analysis of the mobility table);[44] (c) the index does not take into

TABLE 5. COMPARISON OF THREE DIFFERENT KINDS OF INDICES PERTAINING TO EACH STATUS CATEGORY IN BRITISH AND DANISH SAMPLES: (1) INDEX OF PERSISTENCE $\hat{G}_j$, (2) INDEX OF IMMOBILITY BASED UPON THE USUAL MOBILITY RATIO, (3) INDEX OF CONDITIONAL UNCERTAINTY

Table 5A: British Sample

Status Category	Persistence Index	Immobility Index	Uncertainty Index
U	0.40	1.71	0.43
M	−0.22	1.16	0.46
L	0.32	1.67	0.43

Table 5B: Danish Sample

Status Category	Persistence Index	Immobility Index	Uncertainty Index
U	0.52	1.51	0.38
M	−0.21	1.29	0.46
L	0.32	1.97	0.44

[41] See earlier discussion of Table 1B. In addition to the differences between the two indices noted above, we also find that, although the persistence index for the U status category is larger than for the L status category, the usual index applied to the Danish data would have conveyed the impression that the opposite was the case.

[42] This index is based upon a mathematical theory of "information" developed in a different context by Shannon (see, e.g., Shannon and Weaver, 1949). We shall not describe the calculation of the uncertainty index here, but shall instead refer the reader to the details in McFarland's article (1969). To avoid any unnecessary confusion, we calculated this index using logarithms to the base 10, as was done also by McFarland.

[43] In other words, this index does not distinguish between the diagonal cells and the non-diagonal cells. Note that the usual immobility index does distinguish between these cells since it uses the mobility ratio in the diagonal cells to form the index. The index of persistence introduced herein also distinguishes between these cells.

[44] The two distributions may differ from each other in important respects, but their index value will be the same if each of them is different to the same extent (in a certain special sense) from the two particular kinds of extreme distributions described above.

account the magnitudes of the column marginal totals (i.e., the distribution of the destination status categories for the individuals summarized in the mobility table).[45] Because of (a), (b), and (c), we would not be surprised to find situations where the use of the "uncertainty" index led to an impression of a given mobility table that was incompatible with the interpretation obtained with our persistence index.

To compare the two indices, let us reexamine Table 1B. Recall that, aside from the excess of people in the (U,U) cell, there is "perfect mobility" in this table; i.e., (a) there is status persistence of origin status U, and (b) there is no status persistence of origin status M or L. But the numerical values of the "uncertainty" index for this table are 0.47, 0.44, and 0.44, for the U, M, and L status categories respectively; indicating that the distribution of destination status categories for those with origin status U resembles the *uniform* distribution more closely than does the corresponding distributions for those with origin status M or L. This index took into account neither (a) the fact that the distribution of the $\hat{R}_j$ (for $j=1,2,3$) is very different from the uniform distribution in this case (see earlier section herein), nor (b) the fact that the distribution of the column marginals (i.e., the distribution of the destination status categories for the individual summarized in the mobility table) is also very different from the uniform distribution.

Thus, applying the "uncertainty" index to Table 1B, we found that the destination status category is more "uncertain" (in the special sense described above) for an individual whose origin is status U than it is for an individual whose origin status is M or L; but we also had noted earlier that there is positive status persistence of origin status U, and no status persistence of origin status M or L. These comparisons should help to clarify what these indices are actually measuring.[46]

Before closing this section, we comment briefly on three other indices introduced in my earlier article (1969): (1) a modified form of the usual index of status immobility obtained by replacing the usual "expected frequencies" by the entries obtained in the "modified table" in which the expected frequencies are calculated under quasi-perfect mobility; (2) an index based upon the interactions pertaining to "intrinsic status inheritance" of the various status categories;[47] (3) an index based upon the difference between the observed frequencies and the entries obtained in the "modified table" (in which the expected frequencies are calculated under quasi-perfect mobility) relative to the total number of individuals summarized in the mobility table.[48] Although these indices are different from the index $\hat{G}_j$ introduced herein, they all make use of the quasi-perfect mobility model. The index $\hat{G}_j$ is asymmetric with respect to the row and column classifications of the table (it describes persistence *from* origin status *to* destination status); whereas the indices in my earlier articles are symmetric. Applied to the British and Danish samples, the interpretation of the indices in my earlier article was compatible with the present interpretation using $\hat{G}_j$.

COMPARISON OF MAGNITUDES OF THE INDEX OF PERSISTENCE

Having calculated the index of persistence $\hat{G}_j$ for the British and Danish samples in

---

[45] The usual immobility index takes into account the magnitudes of both the column marginals and the row marginals; and the index of persistence introduced herein also takes into account (in a somewhat different way) these marginals calculated for the table with some of its entries blanked out.

[46] For additional comparisons, see Pullum, 1970.

[47] The anti-logarithm of the interaction is a convenient index for reasons of the kind discussed on p. 34 of my earlier article (1969). See the discussion there of the relationship between the new index of immobility and a certain kind of geometric average. (In the case considered there, the geometric average was the anti-logarithm of the interaction.)

[48] The relationship between this index and the index $\hat{G}_j$, which we introduced herein, will be described in the section after the following one. Because of this relationship, the term "persistence" was also used on p. 33 of my earlier article (1969) in the discussion of the index introduced there.

an earlier section herein, we found, for example, that the magnitude of the index was larger for status category U than for status category L, that it was larger for status category L than for status category M, and that the index was negative for status category M. Are these observed differences in magnitude statistically significant? We shall now see how to test the null hypothesis that the differences are not statistically significant.

We noted earlier herein that the null hypothesis of quasi-perfect mobility in the non-diagonal cells could be described by equation (4), or, equivalently, by equation (9) or equation (11). To test this null hypothesis, we presented formula (6) for calculating the corresponding chi-square goodness-of-fit statistic; and we obtained a chi-square of 0.6 for the British sample, and a chi-square of 0.8 for the Danish sample, each chi-square statistic having one degree of freedom. We shall now show how to compare these values with the corresponding chi-square goodness-of-fit values obtained when the null hypothesis to be tested is that the magnitude of the persistence G_j is the same for $j=1,2,3$; i.e., that $G_j=G$ for $j=1,2,3$.[49]

When $G_j=G$ for $j=1,2,3$, then equation (11) can be replaced by

$$\pi_{ij}=\begin{cases} G+(1-G)R_i, & \text{for } i=j \\ (1-G)R_j, & \text{for } i\neq j. \end{cases} \quad (14)$$

In my earlier article (1964b), I showed how to test the null hypothesis described by equation (14). This was done by (a) calculating the maximum-likelihood estimate $\hat{\pi}_{ij}$ of π_{ij}, under the null hypothesis;[50] (b) calculating the estimate $\hat{F}_{ij}$ of the "expected frequency" in cell (i,j) of the mobility table (under the null hypothesis), using the formula

$$\hat{F}_{ij}=f_{i.}\hat{\pi}_{ij}, \quad (15)$$

where $f_{i.}$ is the number of observed individuals having origin status i; and then (c) comparing the observed frequency f_{ij} with the corresponding "expected frequency" $\hat{F}_{ij}$ using the chi-square goodness-of-fit statistic

$$X^2=\sum_{i=1}^{3}\sum_{j=1}^{3}(f_{ij}-\hat{F}_{ij})^2/\hat{F}_{ij}, \quad (16)$$

with a $3 \times 1=3$ degrees of freedom.[51] This goodness-of-fit statistic can be used to test the null hypothesis described by equation (14).

Applying this method to the data in Table 2, we obtain chi-square values of 137.3 and 135.3, respectively, for the British and Danish samples. (See Table 6 for the corresponding values of the $\hat{F}_{ij}$.) Thus, the null hypothesis described by equation (14) would be rejected for both the British and Danish data.

TABLE 6. EXPECTED MOBILITY PATTERN FOR BRITISH AND DANISH MALE SAMPLES WHEN IT IS ASSUMED THAT THE MAGNITUDE OF STATUS PERSISTENCE IN BRITAIN IS THE SAME FROM EACH OF THE THREE STATUS CATEGORIES OF ORIGIN, AND A SIMILAR ASSUMPTION IS MADE ABOUT THE MAGNITUDE OF STATUS PERSISTENCE IN DENMARK

Table 6A: British Sample

		Subject's Status		
		U	M	L
Father's Status	U	507.5	369.4	265.1
	M	321.6	837.9	350.5
	L	180.0	273.3	391.7

Table 6B: Danish Sample

		Subject's Status		
		U	M	L
Father's Status	U	603.6	294.7	184.7
	M	210.9	434.4	132.7
	L	143.7	144.2	242.1

The null hypothesis described by equation (14) differs from the null hypothesis of quasi-perfect mobility for the non-diagonal cells only in that the G_j are assumed equal to each other (for $j=1,2,3$) under the null hypothesis described by equation (14) whereas under quasi-perfect mobility the G_j

[49] The second and third sentence of footnote 21 apply here too.

[50] The calculation of the $\hat{\pi}_{ij}$ is described (in summary form) in Appendix A2 herein.

[51] The f_{ij} are the nine entries in the 3 x 3 mobility table (before any of the entries have been blanked out); and, similarly, the $f_{i.}$ used in formula (15) are the row marginal totals for this mobility table. See my earlier article (1964b), and Appendix A2 herein, for further details.

need not be equal to each other (see equation (11)). We noted earlier that the null hypothesis of quasi-perfect mobility was accepted for both the British and Danish data. Under the assumption that there is quasi-perfect mobility in the non-diagonal cells, we can test the null hypothesis that $G_j = G$ (for $j=1,2,3$), by calculating the difference between the statistic (16) (with its 3 degrees of freedom) and the corresponding goodness-of-fit statistic (6) (with its 1 degree of freedom) obtained in testing for quasi-perfect mobility. This difference will have an asymptotic chi-square distribution with $3-1=2$ degrees of freedom, under the null hypothesis that $G_j = G$ (for $j=1,2,3$).

Applying this method to the data in Table 2, we obtain chi-square values of $137.3-0.6 \doteq 136.7$ and $135.3-0.8 \doteq 134.4$, respectively, for the British and Danish samples. Thus, the null hypothesis that $G_j = G$ (for $j=1,2,3$) would be rejected for both the British and Danish data.

Before closing this section, it is perhaps worth noting that, instead of calculating the chi-square goodness-of-fit statistics, we could have calculated the corresponding chi-square likelihood-ratio statistics (see, e.g., Goodman, 1965, 1968). When this is done, the difference between the chi-square likelihood-ratio statistic calculated to test the null hypothesis described by equation (14) and the corresponding chi-square likelihood-ratio statistic calculated to test the null hypothesis of quasi-perfect mobility can be interpreted as the chi-square likelihood-ratio statistic obtained by comparing the likelihood of obtaining the observed data when there is quasi-perfect mobility with the corresponding likelihood of obtaining the observed data when equation (14) is true. Since the findings obtained when these calculations are applied to the British and Danish data are very similar to those presented above, we shall not go into these details here.

AN INDEX OF THE NET AMOUNT OF PERSISTENCE

For the British 3 x 3 mobility table, the numerical values of the index of persistence pertaining to the three status categories (i.e., 0.40, —0.22, 0.32) provide a summary of this aspect of social mobility, and a similar kind of summary is provided by the corresponding three numerical values calculated for the Danish 3 x 3 mobility table. For each of these mobility tables, we shall now show how to obtain a single numerical value that can serve as a measure of the net amount of status persistence in the table.

We noted in an earlier section herein that, when the index of persistence $\hat{G}_j$ is positive, it can be interpreted as the estimated proportion of "stayers" among those having origin status j; and when $\hat{G}_j$ is negative, the index $-\hat{G}_j$ can be interpreted as the estimated proportion of those in origin status j who will be "removed" from this status category after the first stage. Now let us consider a weighted average $\hat{G}$ of the $\hat{G}_j$, using as the weights the proportion of individuals in the mobility table who are in the corresponding origin status categories. In other words, let

$$\hat{G} = \sum_{i=1}^{3} p_{i\cdot} \hat{G}_i, \tag{17}$$

where $p_{i\cdot}$ is the proportion of individuals in the i^{th} row of the mobility table. With the earlier interpretation of the $\hat{G}_j$, we now find from formula (17) that $\hat{G}$ can be interpreted as the estimated proportion of "stayers" in the population minus the estimated proportion of those who would be "removed" from their status category after the first stage. Calculating this index for the British and Danish samples, we obtain $\hat{G}=.11$ and $\hat{G}=.24$, respectively.

The index described by formula (17) can be applied when the magnitudes of status persistence differ greatly for the different status categories of origin, as well as when these magnitudes are similar. (This index can also be used to estimate the parameter G in equation (14) when the null hypothesis described by (14) is assumed to be true; but the statistic θ calculated in Appendix A2 is actually a better estimate of the parameter in (14) in this special case.)

We shall now comment briefly upon an

index which was introduced in my earlier article (1969); viz.,

$$\hat{G}' = \sum_{i=1}^{3} (f_{ii} - \hat{F}'_{ij})/n, \tag{18}$$

where n is the total number of individuals summarized in the mobility table, f_{ii} is the observed frequency in cell (i,i), and $\hat{F}'_{ii}$ is the corresponding entry in the "modified table"[52] in which the expected frequencies are estimated under quasi-perfect mobility, i.e.,

$$\hat{F}'_{ii} = n^\circ \hat{P}_i \hat{R}_i / [1 - \sum_{k=1}^{3} \hat{P}_k \hat{R}_k], \tag{19}$$

where n° is the total number of individuals in the non-diagonal cells of the table. (The F'_{ii} can also be calculated by either one of the following alternative formulae:

$$\hat{F}'_{ii} = f^\circ_{i\cdot}\ \hat{R}_i/(1-\hat{R}_i),$$

or (20)

$$\hat{F}'_{ii} = f^\circ_{\cdot i}\ \hat{P}_i/(1-\hat{P}_i),$$

where $f^\circ_{i\cdot}$ and $f^\circ_{\cdot i}$ are the marginals for the i^{th} row and i^{th} column respectively, in the mobility table with the diagonal entries blanked out.) We find that $\hat{G}' = \hat{G}$. Thus, formula (18) provides an alternative to formula (17) for the calculation of the index $\hat{G}$; and the interpretation of $\hat{G}$ presented herein applies also to the index (18) introduced in my earlier article.

We also find that

$$p_{i\cdot}\ \hat{G}_i = (f_{ii} - \hat{F}'_{ii})/n. \tag{21}$$

Denoting this quantity as $\hat{S}_i$, we see that $\hat{S}_i$ can be interpreted as the estimated proportion of individuals in the total population who are "stayers" in status category i (when $\hat{S}_i > 0$), and that $-\hat{S}_i$ can be interpreted as the estimated proportion of individuals in the total population who would be "removed" from status category i after the first stage (when $\hat{S}_i < 0$). Calculating $\hat{S}_i$ from the data in Table 2, we obtain 0.13, —0.10, and 0.08, for the U, M, and L status categories, respectively, in Britain; and 0.23, —0.07, and 0.07, respectively, in Denmark.

[52] See, for example, Table 17 in my earlier article (1969).

The index $\hat{G}$ defined by formula (17) made use of a particular set of weights in calculating the weighted average. Other sets of weights might also be of interest. For example, we might also be interested in the index

$$\hat{G}^* = \left[\sum_{i=1}^{3} p_{i\cdot}(1-\hat{R}_i)\hat{G}_i \right] / \sum_{i=1}^{3} p_{i\cdot}(1-\hat{R}_i) \tag{22}$$

$$= \left[\sum_{i=1}^{3} (p_{ii} - p_{i\cdot}\hat{R}_i) \right] / \left[1 - \sum_{i=1}^{3} p_{i\cdot}\hat{R}_i \right],$$

where p_{ii} is the proportion of individuals in the mobility table who have both origin status i and destination status i. Calculating $\hat{G}^*$ for the British and Danish samples, we obtain $\hat{G}^* = .20$ and $\hat{G}^* = .30$, respectively. This index is related to, but different from, several indices that had been introduced earlier for measuring "reliability," "consistency," or "concordance."[53]

THE ANALYSIS OF CROSS-CLASSIFICATION TABLES OTHER THAN THE 3 x 3 TABLES

The methods presented in the preceding sections for the analysis of 3 x 3 cross-classification tables can be extended in a direct fashion to provide methods for the analysis of other tables. We shall illustrate the application of some of the extended methods by applying them to the 5 x 5 table (Table 7) obtained by dividing the U status category of the 3 x 3 table into two status categories, and by dividing the L status category into two status categories as well.[54] The British data are given in sufficient detail (and the sample is large enough) to make possible this particular 5 x 5 cross-

[53] See, e.g., Gini's indices described in Goodman and Kruskal (1959), and the consistency index in Suzuki and Takahasi (1968) and in Suzuki (1968). These indices use, as a standard for comparisons, the model in which the row and column classifications are independent; whereas, the more realistic model of "quasi-independence" is used herein, in one form or another. As we have noted earlier, the indices proposed herein (indices $\hat{G}$ and $\hat{G}^*$, in particular) can also be applied to cross-classifications in contexts other than the analysis of mobility.

[54] This particular 5 x 5 table was also considered in my earlier articles (1965, 1969).

TABLE 7. CROSS-CLASSIFICATION OF BRITISH MALE SAMPLE ACCORDING TO EACH SUBJECT'S STATUS CATEGORY (CATEGORY OF DESTINATION) AND HIS FATHER'S STATUS CATEGORY (CATEGORY OF ORIGIN)

		Subject's Status				
		1	2	3	4	5
Father's Status	1	297	92	172	37	26
	2	89	110	223	64	32
	3	164	185	714	258	189
	4	25	40	179	143	71
	5	17	32	141	91	106

classification; but this could not be done with the Danish data. Even for the British data, it was not possible to divide the M status category into two subcategories.

With the finer division into five status categories in Table 7, we would expect that "status persistence" (in a more general sense) might affect not only the number of individuals who are in the same status category as their fathers but also the number of individuals who are in the status category immediately adjacent to their fathers; and in this case we would test the null hypothesis of "quasi-perfect mobility" blanking out those individuals in Table 7 whose origin status and destination status differ by at most one status category. Thus, if we blank out the entries in the corresponding 13 cells in Table 7 (i.e., the 5 cells on the (main) diagonal and the 8 cells on the adjacent subdiagonals), a chi-square goodness-of-fit value of 1.31 is obtained for testing the null hypothesis of "quasi-perfect mobility" (for the individuals who have not been blanked out). Since 13 cells were blanked out in the 5 x 5 table, the chi-square value has $4 \times 4 - 13 = 3$ degrees of freedom (see Goodman, 1965). Thus, we see that the observed pattern of frequencies in the cells that were not blanked out was congruent with the thesis that there was quasi-perfect mobility in those cells.

From the cells in Table 7 that were not blanked out, the "theoretical tendencies" R_j (for $j=1,2,\cdots,5$) can be estimated by the methods given in my earlier articles (1963, 1964a, 1965). Applying these methods to the data in Table 7 (blanking out those individuals whose origin status and destination status differ by at most one status category), we obtain the following estimates: $\hat{R}_1=0.07$, $\hat{R}_2=0.13$, $\hat{R}_3=0.58$, $\hat{R}_4=0.14$, $\hat{R}_5=0.08$. The corresponding values of the index of persistence $\hat{G}_j$ (for $j=1,2,\cdots,5$) are given in the second column of Table 8.

In view of the fact that Table 7 was obtained by a division of the U and L status categories of Table 2 into status categories 1 and 2 and status categories 4 and 5, respectively, of Table 7, we might also expect that it might only be necessary to blank out the four cells in the upper left corner (i.e., the cells (1,1), (1,2), (2,1), (2,2)), the four cells in the lower right corner (i.e., the cells (4,4), (4,5), (5,4), (5,5)), and the cell in the middle (i.e., the cell (3,3)). Indeed, a chi-square goodness-of-fit value of 7.86 is obtained (with $4 \times 4 - 9 = 7$ degrees of freedom) for testing the null hypothesis of quasi-perfect mobility when we blank out the nine cells listed above. Thus, we see that the data in the cells that were not blanked out are congruent with the thesis that there was quasi-perfect mobility in those cells.

TABLE 8. AN INDEX OF THE DEGREE TO WHICH AN INDIVIDUAL'S STATUS OF ORIGIN PERSISTS FROM HIS ORIGIN TO HIS DESTINATION, CALCULATED FOR EACH STATUS CATEGORY OF ORIGIN IN THE BRITISH SAMPLE, USING TWO DIFFERENT METHODS: (A) BLANKING OUT THE THIRTEEN CELLS ON THE MAIN DIAGONAL AND ADJACENT SUBDIAGONALS, AND (B) BLANKING OUT THE NINE CELLS CONSISTING OF THE FIVE DIAGONAL CELLS AND THE CELLS (1,2), (2,1), (4,5), (5,4)

	Index of Persistence	
Status Category	Calculated by Method A	Calculated by Method B
1	0.44	0.43
2	0.10	0.12
3	—0.27	—0.22
4	0.20	0.20
5	0.21	0.20

Earlier we blanked out 13 cells and now we blank out only 9 cells. Thought must be given, in each particular analysis, to determining which cells should be blanked out. Blanking out the 9 cells has the advantage that it blanks out fewer observations, while blanking out the 13 cells has the advantage that it leads to a chi-square goodness-of-fit value that is somewhat smaller in relative terms. (In comparing the observed chi-square values, the difference in their degrees of freedom can be taken into account by comparing the corresponding percentiles pertaining to the observed values.) Let us now compare the results we obtained earlier in the case where 13 cells were blanked out with the corresponding results obtained when only 9 cells are blanked out.

When the 9 cells are blanked out in Table 7, we obtain the following estimates of the "theoretical tendencies" R_j (for $j=1,\cdots,5$): $\hat{R}_1=0.09$, $\hat{R}_2=0.11$, $\hat{R}_3=0.57$, $\hat{R}_4=0.14$, $\hat{R}_5=0.10$. The corresponding values of the index of persistence $\hat{G}_j$ (for $j=1,2,\cdots,5$) are given in the third column of Table 8. We see from Table 8 that the numerical results obtained are quite similar for the two methods of blanking out cells considered herein. It should also be noted that, if an inappropriate set of cells are blanked out in the 5 x 5 cross-classification table (e.g., if only the five diagonal cells had been blanked out in Table 7), then quite different (and misleading) results would have been obtained. (The inappropriateness of blanking out only the diagonal cells is a consequence of the finer division into five status categories in Table 7;[55] there could not be quasi-perfect mobility in the non-diagonal cells of this table.[56])

In this section we have calculated the index of persistence $\hat{G}_j$ from equation (7), where now $j=1,2,\cdots,5$, and where the $\hat{R}_j$ are calculated with a given set of cells blanked out (the diagonal cells together with some given non-diagonal cells). For a given origin status i, if the only destination status category that is blanked out is $j=i$, then the interpretation of $\hat{G}_i$ presented in the earlier section (in terms of the mover-stayer model and the two-stage model) can be applied here too. On the other hand, in considering a given origin status i, if the destination status categories that are blanked out are $j=i$ and also some other values of j, then a modified form of $\hat{G}_i$ might also be considered; viz.,

$$\hat{D}_i=(\hat{B}_i-\hat{T}_i)/(1-\hat{T}_i), \qquad (23)$$

where $\hat{B}_i$ is the observed proportion of individuals in the destination status categories that are blanked out among those whose origin status was i, and $\hat{T}_i$ is the sum of the $\hat{R}_j$ for the values of j corresponding to the destination status categories that are blanked out. The index $\hat{D}_i$ can be interpreted (by a direct extension of the earlier discussion herein) as the proportion of "stayers" among those with origin status i, where "stayers" now include all individuals who are certain to be in one of the blanked out destination status categories.[57] Although the index $\hat{G}_i$ calculated from equation (7) does not have a similar interpretation (in terms of the mover-stayer model and the two-stage model) except when the only destination status category that is blanked out is $j=i$, this index is still meaningful as a measure of the difference between $\hat{A}_i$ and $\hat{R}_i$, relative to the difference $1-\hat{R}_i$.

[55] Recall the remark earlier in this section pertaining to "status persistence" (in a more general sense) from an origin status category to the same or *adjacent* destination status categories. Other ways of viewing this table also confirm the inappropriateness of blanking out only the diagonal cells here.

[56] This can be verified by an examination, by sight alone, of Table 7. As I noted in my earlier article (1965), the model of "quasi-perfect mobility" for a given set of cells (e.g., the non-diagonal cells) implies, among other things, that there is "perfect mobility" for the population in each rectangular (or square) subtable that can be formed from this set of cells. Thus, "quasi-perfect mobility" for the non-diagonal cells of Table 7 would imply that there would be "perfect mobility" for the population in, say, the 2 x 2 subtable corresponding to cells (1,2), (1,4), (5,2), (5,4) of Table 7, the 2 x 2 subtable corresponding to cells (2,1), (2,5), (4,1), (4,5) of the table, etc.; but this implication is obviously contradicted by the data. For further discussion of this method of analysis, see my earlier article (1965); for related matters, see also McFarland, 1968, and Pullum, 1970.

[57] If $\hat{D}_i$ is negative, then an interpretation of $-\hat{D}_i$ can be obtained by a direct extension of the earlier discussion of the two-stage model (rather than the mover-stayer model). A remark analogous to footnote 30 would also be appropriate here.

APPENDIX

A1. How to Test for Quasi-Perfect Mobility in the Non-Diagonal Cells.

Here we shall describe (in summary form) the calculation of the quantity $\hat{F}^{\circ}_{ij}$ used in formula (6); and the quantities $\hat{R}_j$ and $\hat{P}_i$ used at various points in the present article.

Let us consider a K x K mobility table. (For each of the cross-classifications in Tables 1 and 2, $K=3$; and for the cross-classification in Table 7, $K=5$.) First, replace the entries in the K diagonal cells of the mobility table by zeros; and, for the table thus obtained, let $f^{\circ}_{i.}$ and $f^{\circ}_{.j}$ denote the marginal totals for the i^{th} row ($i=1,2,\cdots,K$) and j^{th} column ($j=1,2,\cdots,K$), respectively. Next, we shall calculate two sets of quantities, $(U_1, U_2, \cdots, U_K)$ and $(V_1, V_2, \cdots, V_K)$, which will be used later in calculating the $\hat{R}_j$, $\hat{P}_i$, and $\hat{F}^{\circ}_{ij}$. We calculate the U_i (for $i=1,2,\cdots,K$) and V_j (for $j=1,2,\cdots,K$) by the following iterative procedure:

As the first step in the iterative procedure, take

$$U^0_i=f^{\circ}_{i.} \text{ (for } i=1,2,\cdots,K). \qquad (24)$$

As the $2m^{th}$ step ($m=1,2,\cdots$), take

$$V_j^{2m-1}=f^{\circ}_{.j}/[U_.^{2m-2}-U_j^{2m-2}] \text{ (for } j=1,2,\cdots K), \qquad (25)$$

where $U_.^{2m}=\sum_{i=1}^{K} U_i^{2m}$. As the $(2m+1)^{th}$ step ($m=1,2,\cdots$), take

$$U_i^{2m}=f^{\circ}_{i.}/[V_.^{2m-1}-V_i^{2m-1}] \text{ (for } i=1,2,\cdots K), \qquad (26)$$

where $V_.^{2m-1}=\sum_{j=1}^{K} V_j^{2m-1}$. The iterative steps are continued for $m=1,2,\cdots$, until the desired accuracy is obtained. Then the quantities $\hat{P}_i$, $\hat{R}_j$, and $\hat{F}^{\circ}_{ij}$ are calculated as follows:

$$\hat{P}_i=U_i/\sum_{k=1}^{K} U_k \text{ (for } i=1,2,\cdots,K) \qquad (27)$$

$$\hat{R}_j=V_j/\sum_{k=1}^{K} V_k \text{ (for } j=1,2,\cdots,K) \qquad (28)$$

$$\hat{F}^{\circ}_{ij}=U_iV_j \text{ (for } i\neq j), \qquad (29)$$

where U_i and V_i denote the quantities obtained when the iterative procedure described above has been completed. The researcher who wishes to apply these methods can use the numerical results pertaining to Tables 1A, 1B, 2A, and 2B, which we presented earlier in this article, to check whether he is carrying out these methods correctly.

The methods described above are suitable when the diagonal cells have been blanked out, and "quasi-perfect mobility" (i.e., "quasi-independence") in the non-diagonal cells is to be assumed or tested. These methods can be directly extended to the case where some other specified set of cells have been blanked out, and "quasi-independence" in the cells that have not been blanked out is to be assumed or tested; and it can also be extended to the analysis of rectangular (rather than square) cross-classification tables (i.e., R x C cross-classification tables, where the number R of rows need not be equal to the number C of columns). The details appear in my earlier articles (1963, 1964a, 1968).[58]

A2. How to Test the Hypothesis that the Magnitude of the Index of Persistence G_j is the Same for the Status Categories $j=1,2,\cdots,K$.

Here we shall describe (in summary form) the calculation of the quantity $\hat{\pi}_{ij}$, which is used in formula (15) to calculate $\hat{F}_{ij}$, which is used in turn in formula (16).

Let us again consider a K x K cross-classification table, where f_{ij} denotes the ob-

[58] The method described in my 1968 article uses a starting point for the iterative procedure that will, in some cases, be slightly different from the starting point used in my earlier articles. With the starting point for the iterations presented in my 1968 article, it was easier to prove that the values of $\hat{P}_i$ and $\hat{R}_j$, which are calculated from the quantities obtained by the iterative procedure, will converge to the appropriate numerical estimates (when the entries that are not blanked out in the table are all positive). Nevertheless, it should be noted that, with the starting point for the iterations presented in my earlier articles, it is also possible to prove that the iterative procedure has this desirable feature (see Haberman, 1969). Criteria for deciding when the iterative procedure has been completed to the desired accuracy can be based upon the numerical values of the $\hat{P}_i$ and/or $\hat{R}_j$ (or the $\hat{F}^{\circ}_{ij}$) calculated at various points in the procedure.

served frequency in cell (i,j) of the table. First, we shall calculate a set of K quantities, $(y_1, y_2, \cdots, y_K)$, which will be used later in calculating the $\hat{\pi}_{ij}$. The y_j (for $j=1,2,\cdots,K$) are obtained as the solution of the following set of K linear equations:

$$y_j c_j + \sum_{i\neq j} y_i(b_j - a_i) = b_j \qquad (30)$$
$$(\text{for } j=1,2,\cdots,K),$$

where the constants a_i, b_j, and c_j in these equations are defined as follows:

$$a_i = f^2_{i.}/f_{ii}, \qquad (31)$$
$$b_j = \sum_{i\neq j} a_i, \qquad (32)$$
$$c_j = b_j - a_j + \sum_{i=1}^{K} f^2_{i.}/f_{ij}, \qquad (33)$$

and the symbol $\sum_{i\neq j}$ denotes a summation over all values of i that are different from j; i.e., over all values of i $(i=1,2,\cdots,K)$ except $i=j$. After solving the set of K linear equations (30) by the usual methods to obtain the values y_j (for $j=1,2,\cdots,K$), we then use these values in the first step of the following iterative procedure:

At the first step, take

$$x_j^{(1)} = y_j \ (\text{for } j=1,2,\cdots,K) \qquad (34)$$

and

$$\theta^{(1)} = 1 - \sum_{j=1}^{K} y_j. \qquad (35)$$

At the m^{th} step $(m=2,3,\cdots)$, take

$$x_j^{(m)} = \{s_j\theta^{(m-1)} - [x_j^{(m-1)}]^2\}/[\theta^{(m-1)} - p_{.j}], \qquad (36)$$

and

$$\theta^{(m)} = 1 - \sum_{j=1}^{K} x_j^{(m)}, \qquad (37)$$

with the constants s_j and $p_{.j}$ defined by

$$s_j = (f_{.j} - f_{jj})/n \qquad (38)$$

and

$$p_{.j} = f_{.j}/n, \qquad (39)$$

where $f_{.j}$ is the marginal total in the j^{th} column, and n is the total number of observations summarized in the mobility table. The iterative steps are continued for $m=2,3,\cdots$, until the desired accuracy is obtained. Let θ and x_j denote the quantities obtained when the iterative procedure has been completed. Then the quantity $\hat{\pi}_{ij}$, which we use in formula (15), is calculated as follows:

$$\hat{\pi}_{ij} = \begin{cases} \theta + x_i & \text{for } j=i \\ x_j & \text{for } j\neq i. \end{cases} \qquad (40)$$

The researcher who wishes to apply these methods can use the numerical results pertaining to Tables 2A and 2B (see, e.g., Tables 6A and 6B), which we presented earlier herein, to check whether he is carrying out these methods correctly. It may also be worth noting that, although an iterative procedure is necessary in order to calculate the maximum-likelihood estimate $\hat{\pi}_{ij}$ of π_{ij} in the present context, it is also possible to apply statistical theory to obtain some justification (when the sample size n is large) for the use of certain kinds of estimates of π_{ij} that do not require iterative calculations. For example, considering the estimate π_{ij} of π_{ij}, which is obtained by replacing θ and x_j in formula (40) by $\theta^{(1)}$ and $x_j^{(1)}$, respectively (i.e., by the particular values (35) and (34) obtained at the first step before the iterative calculations are made), we find that some of the properties of this estimate of π_{ij} will be similar to those of the maximum-likelihood estimate when the sample size is large. For further details, see my earlier article (1964b).

REFERENCES

Blumen, I., M. Kogan and P. J. McCarthy.
1955 The Industrial Mobility of Labor as a Probability Process. Cornell Studies of Industrial and Labor Relations, Vol. 4. Ithaca, New York: Cornell University.

Carlsson, G.
1958 Mobility and Class Structure. Lund, Sweden: Gleerup.

Duncan, O. D.
1965 "Methodological issues in the analysis of social mobility." Pp. 51–97 in N. J. Smelser and S. M. Lipset, Social Structure and Mobility in Economic Development. Chicago, Illinois: Aldine.

Glass, D. V. (ed.)
1954 Social Mobility in Britain. Glencoe, Illinois: Free Press.

Goodman, L. A.
1961 "Statistical methods for the mover-stayer model." Journal of the American Statistical Association 56:841–868.
1963 "Statistical methods for the preliminary analysis of transaction flows." Econometrica 31:197–208.
1964a "A short computer program for the analysis of transaction flows." Behavioral Science 9:176–186.
1964b "The analysis of persistence in a chain of multiple events." Biometrika 51:405–411.
1965 "On the statistical analysis of mobility data." American Journal of Sociology 70:564–585.

Goodman, L. A.
1968 "The analysis of cross-classified data: Independence, quasi-independence, and interactions in contingency tables with or without missing entries." Journal of the American Statistical Association 63:1091–1131.
1969 "How to ransack social mobility tables and other kinds of cross-classification tables." American Journal of Sociology 75:1–40.
Goodman, L. A. and W. H. Kruskal.
1959 "Measures of association for cross classifications. II: Further discussion and references." Journal of the American Statistical Association 54:123–163.
Haberman, S.
1969 Unpublished paper. Department of Statistics, University of Chicago.
McFarland, D. D.
1968 "An extension of conjoint measurement to test the theory of quasi-perfect mobility." Michigan Studies in Mathematical Sociology, Paper Number 3.
1969 "Measuring the permeability of occupational structures: An information-theoretic approach." American Journal of Sociology 75:41–61.
Pullum, T. W.
1970 "What can mathematical models tell us about occupational mobility." Sociological Inquiry 40 (forthcoming).
Rogoff, N.
1953 Recent Trends in Occupational Mobility. Glencoe, Illinois: Free Press.
Shannon, C. E. and W. Weaver.
1949 The Mathematical Theory of Communication. Urbana, Illinois: University of Illinois Press.
Suzuki, T.
1968 "A study on the reliability of social surveys." Proceedings of the Institute of Statistical Mathematics (Tokyo) 16:47–102.
Suzuki, T. and K. Takahasi.
1968 "On an index of consistency of analyzing the data from multi-wave panels." Proceedings of the Institute of Statistical Mathematics (Tokyo) 16:103–107.
Svalastoga, K.
1959 Prestige, Class and Mobility. London: William Heinemann, Ltd.
White, H. C.
1963 "Cause and effect in social mobility tables." Behavioral Science 7:14–27.
Wilensky, H. L.
1966 "Measures and effects of social mobility." Pp. 98–140 in N. J. Smelser and S. M. Lipset, Social Structure and Mobility in Economic Development. Chicago, Illinois: Aldine.
Yasudo, S.
1964 "A methodological inquiry into social mobility." American Sociological Review 29:16–23.

MEASURING POPULATION DIVERSITY *

STANLEY LIEBERSON
University of Washington

Viewing diversity as the position of a population along a homogeneity-heterogeneity continuum, a general method is presented for describing diversity within and between groups that are classified by one or more qualitative variables. This method has a wide range of applications, including such phenomena as attitudinal concensus, political cleavage, residential isolation, linguistic communication, cohesion, as well as the general diversity of populations. Diversity is operationally defined as the probability of obtaining unlike characteristics when two persons are randomly paired. Computation of the indexes of diversity within a population, A_w, and between two populations, A_b, is illustrated with data drawn from several substantive areas in sociology.

This paper proposes a general method for describing the magnitude of diversity *within* and *between* social aggregates. The method can be applied to populations classified by one or more qualitative variables, for example, religion, ethnic origin, political party, etc. Since diversity measures are appropriate for either attitudinal or social characteristics, their range of application is rather extensive, including such phenomena as political cleavage, cohesion, consensus, and residential isolation. The approach taken, based on the elementary application of permutations and combinations

* I am deeply indebted to Mrs. Lynn K. Hansen for her invaluable assistance on this project. The paper benefited greatly from a critical reading by Professor Herbert L. Costner. This paper is a byproduct of research conducted on language pluralism under N.S.F. Grant GS–1869.