DEMOGRAPHY© Volume 12, Number 4 November 1975

SOME DETERMINANTS OF THE MIGRATION OF PROFESSIONAL MANPOWER

Dieter Wermuth
German Council of Economics Advisors, Wiesbaden, Stresemannring 11, Federal Republic of Germany
Nanny Wermuth
Institut für Medizinische Statistik und Dokumentation, Johannes-Gutenberg-Universität, Mainz, Langenbeck Str. 1, Federal Republic of Germany

Abstract—Determinants of migration of professional manpower are investigated using data from a 1970 survey of immigrants to the United States. From a respondent's stated "intent to stay" in the United States and five other characteristics a six-dimensional contingency table is formed. We find a well-fitting log-linear model for this table. Thus, we establish the importance of selected determinants of migration and present a table of predicted rates of intent to stay in the United States.

Introduction

Migration theory tries to provide explanations for the migration of people between jobs and between places of residence (cities, states, countries). On the aggregate economic level it is primarily differences in employment opportunities and in wage levels that have been found to be good determinants of the direction and the size of migration flows (Böhning, 1971; Bowles, 1970; Comay, 1970; Grubel and Scott, 1966).

On the micro-analytical level it is of interest to analyze the reasons that cause a person to migrate or not to migrate and thus to attach probabilities of migration during a given time period to persons for whom important characteristics are known. In previous studies these characteristics were, for instance, age, level of education, type of job, and number of dependents (Grubel and Scott, 1967).

So far, little is known about an important aspect of migration, namely, about those migrants who after spending a period of time in their new place of residence decide to return to the place they originally came from or to migrate to a third location. It would be useful to know to what extent an immigration flow is counterbalanced by return migration. All calculations of the social costs of "brain drains," for example, are incomplete or exaggerated if no account is taken of this return migration.

One of the few studies that deal with return migration is an article by Comay (1971). He investigated the probability of individual immigrant Canadians in the United States leaving for Canada again. With "return" as the dependent variable he found that the variables "employment sector," "length of employment in the U. S." and "degree level" explained, in that order of importance, the greatest proportion of the variance. For the present study on Western European professional immigrants we included these variables in addition to some others that we expected could have some influence on the return migration decision. (Rather than "length of employment" we chose "length of stay in the United States" because in this way we could include the time that the immigrant had spent as a student in the United States. This is more

important for European immigrants, most of whom do not speak English, than it is for Canadians.)

Comay used the regression method although the dependent variable "return" was a dichotomous variable. Whenever the dependent variable is qualitative, however, it is known that the error variation changes systematically so that the assumption of homoscedasticity is violated (for a proof see Goldberger, 1964, pp. 249–251), the test statistics will hardly follow their assumed distribution, and the predicted values of the dependent variable may lie outside of the interpretable limits (Theil, 1971; Goodman, 1972). Since, in our data not only the dependent variable but also many of the predetermined variables are qualitative, we decided to analyze the data in a contingency table context.

In the following we are attempting to answer the question: What are the determinants of a professional immigrant's stated intention to stay in or to emigrate from the United States? This is not precisely the problem one is interested in because one would like to know what distinguishes the professional immigrant who remains in the United States from the actual return migrant. But it is easy to imagine the difficulties and expense that would arise in attempting to survey two populations separated by perhaps a couple of thousand miles. Such data are generally not available. We therefore had to rely on intentions rather than on actual facts. The main drawback of using intentions is, of course, that they are formed by a number of opinions which may change easily. There is also often a marked difference between what a person wants to do and what he can do. For the time being, however, we have to rely, as everybody else has, on this sort of second best information.

Source of Data

The data were taken from a survey conducted by the National Science Foundation through the facilities of the U. S. Bureau of the Census. The sample consisted of 5,000 persons from all over the world who had immigrated to the United States, who had not yet received full American citizenship at the time the mailing list was set up (early spring, 1970), and who were either engineers, natural or social scientists. It was believed that the sample was representative of a group of about 50,000 such persons (Waksberg, 1971). The total sample of Western Europeans was narrowed down to those for whom it was known that they had received their primary and secondary schooling in their countries of birth. Only for those persons could it be assumed that they had actually spent their formative years in their countries of birth, enabling us to call them Germans, British, Dutch, etc. The few social scientists were excluded. This left 842 persons for whom we had information on their intention to emigrate from the United States. Seven hundred and ninety-six of the 842 persons had no missing responses on the other five variables. This group formed the data basis for the final model.

Definition of the Variables

The NSF questionnaire contained considerable detailed information on motivations, personal characteristics, and job history. Our dependent variable was based on the question: Do you expect to emigrate from the United States? The possible answers and the number of respondents per answer were as follows.

1	no	466
2	yes, within 1 year	22
3	yes, in 1–3 years	23
4	yes, in 4–5 years	4
5	yes, after 5 years	1
6	yes, undecided when	28
7	undecided	298
		842

Only a small number of the respondents selected any one of the answers 2 to 6. Therefore, we grouped the answers into

only two classes: "No" = no intent to return and "Yes" (numbers 2 to 7) = yes, there is the intent, even though of varying intensity, to emigrate from the United States. Of the other information contained in the original set of data we chose the variables listed in Table 1 as possible explanatory variables for the stated intent to remain in the United States.

In the upper half of Table 2 the chi-square values that evaluate the associations between these eleven variables are shown. The number of respondents per two-way table is displayed in the lower half. The variables employment sector (8), level of education (9), education in the U. S. (10) and major field of study (11) have little or no significant association with intent to return (1). And even though mobility (7) seems, when looked at separately, to be an important determinant of intent to return, we decided against including this variable in a multidimensional contingency table. Firstly, the information on mobility is generally less easily obtained than the information on any one of the other five determinants of migration. Secondly, mobility correlates highly with country

TABLE 1.—Frequency Distributions of Explanatory Variables

Variable		Observed Frequency
Variable 1:	intent to return	
	1 = no	466
	2 = yes	376
		842
Variable 2:	country of origin[a]	
	1 = Central Europe	267
	2 = Scandinavia	157
	3 = France	50
	4 = U. K. and Ireland	179
	5 = Italy	50
	6 = Mediterranean countries (except Italy)	139
		842
Variable 3:	American spouse[b]	
	1 = yes	109
	2 = no	733
		842
Variable 4:	length of stay in the U. S. since the time of immigration	
	1 = 1, 2, 3, 4 years	370
	2 = 5, 6, 7, 8 years	285
	3 = 9 and more years	187
		842
Variable 5:	present job in research and development	
	1 = yes (development; information and data processing, or technical writing; research; teaching, instructing, training)	430
	2 = no (advertising, consultation; design, drafting, drawing, graphics; production, construction, operations, maintenance; sales, marketing, technical services; testing; evaluation, inspection; other)	366
		796

Table 1—Continued

Variable 6:	age at the time of immigration[c]	
	1 = younger than 28 years	270
	2 = 28 - 40 years	496
	3 = older than 40 years	76
		842
Variable 7:	mobility	
	1 = the respondent was a resident for a year or more in 0 or 1 foreign country	581
	2 = the respondent was a resident for a year or more in 2 or more foreign countries	258
		839
Variable 8:	employment sector	
	1 = industry	654
	2 = government	11
	3 = other (university or college; other non-profit institution)	117
		782
Variable 9:	level of education	
	1 = less than Bachelor's degree	103
	2 = Bachelor's degree or equivalent	303
	3 = Master's degree or equivalent	250
	4 = Doctorate	185
		841
Variable 10:	education in the U. S.	
	1 = no	553
	2 = yes	289
		842
Variable 11:	major field of study	
	1 = engineering	584
	2 = natural science	203
		787

a- Countries were grouped according to geographical area and common characteristics of their populations such as similar GNP per capita.

b- The citizenship of the spouse at the time of immigration--if there was a spouse--was used as an indicator for the existence or the nonexistence of close personal ties to the U. S.

c- The cutting points 28 and 40 years roughly represent the age in a person's professional life at which he has finished his formal education, and at which he does not move between jobs very frequently.

Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

of origin and age at the time of immigration. Finally, by excluding this one variable we could reduce the number of cells in the contingency table from 864 to 432. Thus, we obtained a better relationship between the number of respondents (i.e., the number of observations) and the number of parameters to be estimated.

Model Search

The main purpose of analyzing the six-dimensional contingency table was to predict intent to stay from known characteristics of an immigrant.

One natural approach would have been to fit a logit model to the observed rates of intent to stay as described, for instance, by Dyke and Patterson (1952). However, because of the large number of zero cell counts in our data table (Table 3), it was not feasible to compute from the observed counts the ratio of intent

TABLE 2.—Chi-Square Values and Significance Levels for Two-Way Tables (Upper Half) and Number of Respondents per Table (Lower Half)

Variable	Degrees of Freedom	Variable 1	2	3	4	5	6	7	8	9	10	11
		Degrees of Freedom										
		1	5	1	2	1	2	1	2	3	1	1
1	1	X	39.23[b]	14.87[b]	9.37[b]	11.05[b]	16.48[b]	7.60[b]	3.47	9.83[a]	3.63	4.17[a]
2	5	842	X	68.22[b]	26.34[b]	12.42[a]	5.01	20.95[b]	26.85[b]	93.15[b]	122.10[b]	38.81[b]
3	1	842	842	X	4.05	0.02	3.01	5.48[a]	6.26[a]	6.0	33.05[b]	0.32
4	2	842	842	842	X	2.95	6.67	4.05	5.65	40.33[b]	3.69	1.02
5	1	796	796	796	796	X	2.95	2.46	90.92[b]	187.80[b]	7.70[b]	112.99[b]
6	2	842	842	842	842	796	X	40.34[b]	6.41	15.01[a]	22.95[b]	4.52
7	1	839	839	839	839	796	839	X	1.41	6.77	8.56[b]	0.32
8	2	782	782	782	782	782	782	782	X	160.83[b]	9.81[b]	91.17[b]
9	3	841	841	841	841	795	841	838	781	X	37.03[b]	218.01[b]
10	1	842	842	842	842	796	842	839	782	841	X	1.83
11	1	787	787	787	787	744	787	784	733	786	787	X

a- Significant at .05 level.
b- Significant at .01 level.

Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

to stay to intent to emigrate in each of the 216 (= 6 × 2 × 3 × 2 × 3) categories. Therefore, we first searched for log-linear models (Birch, 1963; Bishop, 1967, 1969b; Goodman, 1970) for the full six-dimensional contingency table, which described well the interrelations between all variables. Formally, at most 432 parameters can be estimated in a 2 × 6 × 2 × 3 × 2 × 3 table, and many more log-linear models exist (Goodman, 1970). But fortunately the search for a reasonable model required the evaluation of relatively few models. Selection techniques among log-linear models have, for instance, been described by Goodman (1970, 1971a) and Wermuth (1975b).

The first hypothesis that we tested was whether all variable pairs that are marginally independent (see Table 2) can be regarded as conditionally independent as well. The resulting multiplicative model is denoted as 123/124/125/16 (compare the Appendix). It has 368 degrees of freedom and a likelihood-ratio chi-square statistic with value 385.92. This represents a non-significant result. Therefore, the aforementioned hypothesis may be accepted.

An interpretation of the model is that the variables age at time of immigration (6), research and development job (5), length of stay (4), and American spouse (3) are uncorrelated for each group of persons with known intent to return and known country of origin (that is, for each category of the joint variable 12). Furthermore, the variables country of origin (2) and length of stay (6) are not associated for the group of immigrants with intent to stay and for those with intent to return (that is, variables (2) and (6) are conditionally independent given variable (1)).

In a second step, we decided that no simpler multiplicative model fits the data (Wermuth, 1975b). Finally, we found the simplest non-multiplicative model (shown in Table 5) that is consistent with the data (Goodman, 1971a).

More precisely, after we had accepted pattern 123/124/125/16, we determined how many interactions (Birch, 1963) can be assumed to be zero. We could exclude all three-factor interactions and none of the two-factor interactions. Therefore, we accepted a model denoted as 12/13/14/15/16/23/24/25. This model has as minimal sufficient statistics the observed two-way tables of all the listed variable pairs. This is to say that these two-way tables are sufficient to approximately reproduce the observed cell counts in the six-dimensional contingency table.

Results

We found that model 12/13/14/15/16/23/24/25 describes the interrelations among all variables rather well. Thus, the important associations between the variables are those among the eight listed variable pairs. Table 4 permits a close look at these associations.

For Western Europeans the average intent to stay in the United States is 55.4 percent. For persons from the Mediterranean countries (except Italy), from France and from the Central European countries the intent to stay is above average. It is 75.4 percent, 60.0 percent, and 56.5 percent, respectively. With 41.6 percent the Scandinavians show the lowest intent to stay, followed by the Italians (50.0 percent).

The variables American spouse, length of stay, and age at the time of immigration all exert the expected effect on the intent to stay. An American spouse has a strong influence in favor of staying in the United States (71.7 percent versus 52.9 percent). The longer a person has been an immigrant the more likely he is to stay, and the older a person is when he immigrates, the more determined he is to make the United States his permanent home.

One surprising finding is that persons employed in either research, development, teaching, or data processing show

TABLE 3.—Observed Cell Counts in the Six-Dimensional Contingency Table

Age at Time of Immigration	Present Type of Job	Length of Stay	Amer. Spouse	Country of Origin											
				CEur		Scan		Fran		GBrit		Italy		Medit	
				s.	r.	s.	r.	s.	r.	s.	r.	s.	r.	s.	r.
Under 28	R and D	1-4	yes	0	1	0	0	1	0	1	0	0	0	3	0
			no	1	9	0	4	1	1	1	14	1	1	2	0
		5-8	yes	1	1	2	0	1	0	1	0	0	0	1	1
			no	8	4	2	11	3	1	2	5	1	4	2	1
		>8	yes	0	0	0	0	1	0	1	0	0	0	2	0
			no	5	7	3	4	0	0	2	2	0	0	0	0
	Not R and D	1-4	yes	2	1	0	0	0	0	0	0	1	0	5	1
			no	3	6	2	10	1	0	2	7	1	0	7	1
		5-8	yes	2	0	1	0	0	1	0	0	0	1	2	0
			no	9	7	4	6	3	0	4	4	1	0	2	0
		>8	yes	1	0	0	0	1	2	0	0	0	0	1	0
			no	10	1	3	3	0	0	4	2	0	1	0	0
28-40															
	R and D	1-4	yes	0	2	0	0	1	0	2	2	1	0	7	3
			no	14	16	2	15	5	6	7	16	5	5	9	4
		5-8	yes	0	0	2	0	0	0	0	0	0	1	6	0
			no	15	14	8	7	1	3	10	4	0	6	8	5
		>8	yes	0	0	1	2	1	0	0	0	1	2	2	1
			no	10	6	4	5	4	0	6	5	1	0	3	0
	Not R and D	1-4	yes	1	0	0	0	2	1	0	1	2	1	1	0
			no	13	11	9	10	2	1	15	7	1	0	11	6
		5-8	yes	1	0	0	0	0	0	0	0	1	0	0	1
			no	14	9	6	5	0	0	8	4	0	0	3	2
		>8	yes	0	1	1	0	0	1	3	2	1	0	2	0
			no	14	8	6	2	0	0	7	4	1	0	4	1
40 and over															
	R and D	1-4	yes	0	0	0	0	0	0	0	0	0	0	0	0
			no	5	1	1	0	0	0	0	2	1	0	4	0
		5-8	yes	1	0	0	0	0	0	0	0	0	0	0	0
			no	4	2	1	1	0	1	0	0	2	0	1	1
		>8	yes	2	0	0	0	0	0	0	0	0	0	0	0
			no	1	0	1	0	1	1	2	0	0	0	1	0
	Not R and D	1-4	yes	0	0	0	0	0	0	1	0	0	0	0	0
			no	3	2	2	0	1	1	2	1	0	0	3	0
		5-8	yes	1	0	1	0	0	0	0	0	0	0	0	0
			no	0	1	0	1	0	0	6	0	1	0	2	0
		>8	yes	0	0	0	0	0	0	0	0	0	0	0	0
			no	2	0	0	1	0	0	0	1	0	0	0	0

s. = stay
r. = return
R and D = Research and Development
CEur = Central Europe
Scan = Scandinavia
Fran = France
GBrit = U.K. and Ireland
Medit = Mediterranean countries (except Italy)

Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

TABLE 4.—Marginal Sums and Column Percentages for the Variable Pairs (1,2) (1,3) (1,4) (1,5) (1,6) (2,3) (2,4) (2,5)

Marginal table 12:

Intent to Return (1)	Country of Origin (2)						
	CEur	Scan	Fran	GBrit	Italy	Medit	Sum
No	143 (56.5)	62 (41.6)	30 (60.0)	87 (51.2)	24 (50.0)	95 (75.4)	441 (55.4)
Yes	110	87	20	83	24	31	355
Sum	253	149	50	170	48	126	796

Marginal table 13:

Intent to Return (1)	American Spouse (3)		
	Yes	No	Sum
No	76 (71.7)	365 (52.9)	441 (55.4)
Yes	30	325	355
Sum	106	690	796

Marginal table 14:

Intent to Return (1)	Length of Stay in Years (4)			
	1-4	5-8	Over 8	Sum
No	170 (49.3)	155 (57.4)	116 (64.1)	441 (55.4)
Yes	175	115	65	355
Sum	345	270	181	796

a lower intent to stay in this country than do others. Per capita expenses for research and development are considerably higher in the United States than in any of the Western European countries; hence it could be expected that research facilities and employment opportunities have been more favorable than in Europe. But, since the second half of the sixties many European countries have

Table 4—Continued

Marginal table 15:

Intent to Return (1)	Present Type of Job (5)		
	R and D	Not R and D	Sum
No	215 (50.0)	226 (61.7)	441 (55.4)
Yes	215	140	355
Sum	430	366	796

Marginal table 16:

Intent to Return (1)	Age at the Time of Immigration in Years (6)			
	Under 28	28-40	Over 40	Sum
No	123 (48.6)	265 (56.0)	53 (75.7)	441 (55.4)
Yes	130	208	17	355
Sum	253	473	70	796

Marginal table 23:

American Spouse (3)	Country of Origin (2)						
	CEur	Scan	Fran	GBrit	Italy	Medit	Sum
Yes	18 (7.1)	10 (6.7)	13 (26.0)	14 (8.2)	12 (25.0)	39 (31.0)	106 (13.3)
No	235	139	37	156	36	87	690
Sum	253	149	50	170	48	126	796

undertaken considerable efforts to close their technology gap vis-a-vis the United States. Perhaps this actually created a better employment climate for scientists there.

Three other variables are associated with country of origin: immigrants from France, Italy, and the other Mediterranean countries are much more likely to be married to an American and are also more likely to have jobs in research and development than are immigrants from Central and Northern Europe (marginal tables 23 and 25). Further, the flow of immigrants during certain periods differed from country to country

Table 4—Continued

Marginal table 24:

Length of Stay in Years (4)	Country of Origin (2)						
	CEur	Scan	Fran	GBrit	Italy	Medit	Sum
1-4	91 (36.0)	55 (36.9)	24 (48.0)	81 (47.6)	23 (47.9)	71 (56.3)	345 (43.3)
5-8	94 (37.2)	58 (38.9)	14 (28.0)	48 (28.2)	18 (37.5)	38 (30.2)	270 (33.9)
>8	68 (26.8)	36 (24.2)	12 (24.0)	41 (24.1)	7 (14.6)	17 (13.5)	181 (22.7)
Sum	253	149	50	170	48	126	796

Marginal table 25:

Present Type of Job (5)	Country of Origin (2)						
	CEur	Scan	Fran	GBrit	Italy	Medit	Sum
R and D	130 (51.4)	76 (51.0)	33 (66.0)	85 (50.0)	35 (72.9)	71 (56.3)	430 (54.0)
Not R and D	123	73	17	85	13	55	366
Sum	253	149	50	170	48	126	796

Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

(marginal table 24). During the four years before the survey was taken, immigration from Central Europe and Scandinavia was below average. Similarly, below average immigration was observed for France and Great Britain from 1962 to 1967 and for Italy and the other Mediterranean countries before 1962. No truly convincing explanations can be offered for the relationships in these marginal tables.

Table 6 shows predicted rates of intent to stay obtained from the expected cell counts for our selected model. Thus, rather than looking at the direction and the effect of each of the five determinants of migration separately, as in Table 5, we took into account the different influences simultaneously. For each of the 216 categories (of the joint variable 23456) we computed a rate—or a probability to stay—as the ratio of the expected intent to stay to the sum of the expected intent to return and the ex-

TABLE 5.—Likelihood Ratio Tests for Selected Log-Linear Models and for Interactions

Model	Likelihood Ratio Chi-Square	d.f.	Conditional Test for the Interaction Among Variables	Likelihood Ratio Chi-Square	d.f.
123/124/125/16	385.92	368			
13/23/124/125/16	389.47	373	123	3.55	5
13/23/14/24/125/16	402.07	383	124	12.60	10
12/13/14/15/16/23/24/25	406.76	388	125	4.69	5
13/14/15/16/23/24/25	441.30	393	12	34.54	5
12/14/15/16/23/24/25	413.59	389	13	6.83	1
12/13/15/16/23/24/25	423.47	390	14	16.71	2
12/13/14/16/23/24/25	419.81	389	15	13.05	1
12/13/14/15/23/24/25	423.96	390	16	17.20	2
12/13/14/15/16/24/25	456.47	393	23	49.71	5
12/13/14/15/16/23/25	436.79	398	24	30.03	10
12/13/14/15/16/23/24	421.55	393	25	14.79	5

d.f. = degrees of freedom.
Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

pected intent to stay. Clearly, with so many zero cell counts in the six-dimensional table (Table 3), we cannot expect to obtain very precise and stable estimates of the intent to stay in the United States. A much larger sample is necessary to validate our tentative results.

For each country, the lowest probability to stay is observed for respondents who are not married to an American, who immigrated between one to four years before the time of the survey, who hold a job in the research and development field, and who are younger than twenty-eight years. The rates for this group of persons vary between 22 percent for Scandinavians and 56 percent for immigrants from Mediterranean countries. The highest probabilities to stay vary considerably less, from 85 percent to 96 percent. These highest rates hold for persons who are married to an American, who have stayed in the United States for more than eight years already, who do not work in research and development, and who were over the age of forty at the time of immigration. Table 6 also permits the identification of groups of immigrants with about equal probability of staying in the United States.

TABLE 6.—Predicted Rates of Intent to Stay in the United States (in Percentages)

Age at Time of Immigration	Present Type of Job	Length of Stay	Amer. Spouse	Country of Origin					
				CEur	Scan	Fran	GBrit	Italy	Medit
Under 28	R and D	1-4	yes	48	34	53	45	45	70
			no	34	22	38	30	30	56
		5-8	yes	59	44	70	55	55	79
			no	44	30	49	40	40	66
		>8	yes	67	53	71	63	63	84
			no	52	37	56	48	48	73
	Not R and D	1-4	yes	62	46	67	57	58	80
			no	47	32	51	42	42	69
		5-8	yes	72	58	74	68	68	86
			no	58	42	62	53	54	77
		>8	yes	78	65	80	75	75	90
			no	65	50	69	61	61	83
28-40	R and D	1-4	yes	56	41	68	52	52	76
			no	41	27	45	37	37	63
		5-8	yes	66	52	70	62	62	83
			no	52	37	60	47	47	73
		>8	yes	73	60	77	69	69	87
			no	60	44	64	55	55	79
	Not R and D	1-4	yes	69	54	73	64	65	85
			no	54	39	59	50	50	75
		5-8	yes	77	65	80	74	74	89
			no	65	50	68	60	61	82
		>8	yes	82	72	85	80	79	92
			no	72	58	75	68	67	86
Over 40	R and D	1-4	yes	75	82	78	73	74	86
			no	63	48	67	59	59	81
		5-8	yes	84	71	87	78	80	93
			no	72	59	76	69	69	87
		>8	yes	88	82	90	88	85	94
			no	78	66	81	75	75	90
	Not R and D	1-4	yes	83	73	86	81	80	93
			no	74	61	77	71	71	88
		5-8	yes	89	83	93	86	86	95
			no	82	71	85	79	76	92
		>8	yes	93	85	92	90	83	96
			no	86	77	87	84	86	95

s. = stay
r. = return
R and D = Research and Development
CEur = Central Europe
Scan = Scandinavia
Fran = France
GBrit = U.K. and Ireland
Medit = Mediterranean countries (except Italy)

Source: 1970 National Science Foundation Survey of Immigrant Scientists and Engineers data tape.

Appendix

Model 123/124/125/16 may be derived, characterized and interpreted in a number of ways. For instance, the following characterizations are equivalent (Birch, 1963; Bishop, 1969a; Goodman, 1970; Wermuth, 1975a):

The expected cell count m_{ijklrs} can be computed multiplicatively from the marginal tables with variables 123, 124, 125, 16, that is, from $m_{ijk}\ldots = \Sigma_{lrs} m_{ijklrs}$, etc., as

$$m_{ijklrs} = \frac{m_{ijk\ldots}\, m_{ij.l..}\, m_{ij..r.}\, m_{i....s}}{m_{ij....}^{2}\, m_{i.....}}.$$

In the log-linear model for the six variables, all interaction terms involving the variable pairs (3,4), (3,5), (3,6), (4,5), (4,6), (5,6), (2,6) are zero.

All seven of these variable pairs have zero partial associations, that is, each pair of variables is conditionally independent given the remaining four variables.

Of the possible interpretations for the model we describe three. Using the methods and terminology presented in Goodman (1973) we obtain the following:

At each level of variables 1 and 2, the four variables 3, 4, 5, and 6 are mutually independent, *and* the effects in the model are "heterogeneous" with respect to the level of variables 1 and 2 except for the effect of variable 6, which is "homogeneous" with respect to the level of variable 2 and heterogeneous with respect to the level of variable 1.

Using the methods for multiplicative models proposed by Wermuth (1975a) or using the methods for general log-linear models presented by Goodman (1970, 1971b) we arrive at the following two interpretations.

(1) Variable 6 is independent of the joint variable 2345, given the levels of variable 1, *and* variables 3, 4, and 5 are independent, given the levels of the joint variable 12. (The first statement is implied by the zero partial associations for pairs (2,6), 3,6), (4,6), (5.6), which lead to model 16/12345. The second statement follows from the additional zero partial associations for the pairs (3,4), (3,5), and (4,5).)

(2) Variables 3, 4, 5, and 6 are independent, given the levels of the joint variable 12, *and* variables 2 and 6 are independent, given the levels of variable 1. (The first statement is implied by the zero partial associations for pairs (3,4), (3,5), (3,6), (4,5), (4,6), (5,6), which lead to model 123/124/125/126. The second statement follows from the additional zero partial association for pair (2,6).)

Acknowledgments

We thank Paul Holland for stimulating discussions and the referees for their helpful comments. Computer time was provided by the Cambridge Project through the Department of Statistics, Harvard University, and by the Department of Economics, Tufts University.

References

Birch, M. W. 1963. Maximum Likelihood in Three-Way Contingency Tables. Journal of the Royal Statistical Society, Series B 25: 220–233.

Bishop, Yvonne M. M. 1967. Multi-Dimensional Contingency Tables. Unpublished Ph.D. dissertation. Cambridge: Harvard University.

———. 1969a. Calculating Smoothed Contingency Tables. Pp. 273–286 in J. P. Bunker, W. H. Forrest, F. Mosteller, and L. D. Vandam (eds.), The National Halothane Study. Bethesda: National Institute of Health.

———. 1969b. Full Contingency Tables, Logits, and Split Contingency Tables. Biometrics 25: 383–399.

Bohning, W. R. 1970. The Differential Strength of Demand and Wage Factors in Intra-European Labour Mobility: With Special Reference to West Germany, 1957–1968. International Migration 8:193–202.

Bowles, Samuel. 1970. Migration as Investment: Empirical Tests of the Human Investment Approach to Geographic Mobility. Review of Economics and Statistics 52:356–362.

Comay, Yochanan. 1969. International Migration of Professional Manpower: The Canada-U. S. Case. Unpublished Ph.D. dissertation. Princeton: Princeton University.

———. 1971. Determinants of Return Migration: Canadian Professionals in the U. S. Southern Economic Journal 37:318–322.

Dyke, G. V., and H. D. Patterson. 1952. Analysis of Factorial Arrangements When the Data Are Proportions. Biometrics 8:1–12.

Goldberger, Arthur S. 1964. Econometric Theory. New York: John Wiley and Sons.

Goodman, Leo A. 1970. The Multivariate Analysis of Qualitative Data: Interactions among Multiple Classifications. Journal of the American Statistical Association 65:226–256.

———. 1971a. The Analysis of Multidimensional Contingency Tables: Stepwise Procedures and Direct Estimation Methods for Building Models for Multiple Classifications. Technometrics 13:33–61.

———. 1971b. Partitioning of Chi-Square, Analysis of Marginal Contingency Tables, and Estimation of Expected Frequencies in Multidimensional Contingency Tables. Journal of the American Statistical Association 66:339–344.

———. 1972. A Modified Multiple Regression Approach to the Analysis of Dichotomous Variables. American Sociological Review 37: 28–46.

———. 1973. Guided and Unguided Methods for the Selection of Models for a Set of *T* Multidimensional Contingency Tables. Journal of the American Statistical Association 68:165–175.

Grubel, H. G., and A. D. Scott. 1966. The Immigration of Scientists and Engineers to the United States, 1949–61. Journal of Political Economy 74:368–378.

———, and A. D. Scott. 1967. Determinants of Migration: The Highly Skilled. International Migration 5:127–139.

Theil, Henri. 1971. Principles of Econometrics. New York: John Wiley and Sons.

Waksberg, J. 1971. Memorandum on NSF immigrant scientists and engineers. Statistical Methods Division, U. S. Bureau of the Census, July 26.

Wermuth, D. 1972. Economic Determinants of the Migration of Professional Manpower between Western Europe and the United States. Unpublished Ph.D. dissertation. Medford, Massachusetts: Tufts University.

Wermuth, N. 1975a. Analogies between Multiplicative Models in Contingency Tables and Covariance Selection. Biometrics. Forthcoming.

———. 1975b. Model Search among Multiplicative Models. Biometrics. Forthcoming.