FILE COPY
DO NOT REMOVE

#495-78

INSTITUTE FOR RESEARCH ON POVERTY DISCUSSION PAPERS

EARNINGS AND CAPABILITY REQUIREMENTS

Joop Hartog

UNIVERSITY OF WISCONSIN ~ MADISON

Earnings and Capability Requirements

Joop Hartog

May 1978

The author, from the Institute for Economic Research, Erasmus University, Rotterdam is presently visiting Project Associate at the Institute for Research on Poverty, University of Wisconsin-Madison.

The research reported here was supported by funds granted to the Institute for Research on Poverty at the University of Wisconsin-Madison by the Department of Health, Education, and Welfare pursuant to the provisions of the Economic Opportunity Act of 1964. The conclusions expressed herein are those of the author.

ABSTRACT

This paper empirically investigates the relationship between earnings and capability requirements in the United States. Emphasis is on the need to use data on the capability requirements of an individual's job rather than on an individual's capability endowments. Data are taken from the 1950 and 1960 Censuses and from the Dictionary of Occupational Titles (DOT). Using factor analysis, the DOT data are searched for some underlying basic capabilities; next, implicit capability prices are estimated. Non-linearity in the earnings function is analyzed, and price changes over time are studied as well.

Earnings and Capability Requirements

1. THEORY AND MOTIVATION

There is more speculation than solid evidence regarding the role of individual abilities in shaping the personal distribution of earnings. The matter is of obvious political and scientific interest, yet economists have either left most of the work in this area to professions such as occupational psychology and education or have chosen not to deal with it. In most economists' approaches, ability comes in as a variable that might bias the estimate of other coefficients, such as that of schooling, but is not considered essential in its own right.[1]

This paper takes a different approach. The work reported here stems from a theory developed by the author (1978a, Chapter 9; a summary is given in 1978b), which in turn draws its inspiration from Tinbergen (1956) and Roy (1951). In this theory, the labor market is structured in terms of capabilities, those characteristics of an individual that determine his or her productive potential. A job is described by demanded levels of capabilities, an individual by supplied levels. Individuals are supposed to command stocks of capabilities and to decide on their job choice by setting the desired level of supply of each of their capabilities. This implies the emergence of an important distinction between capability endowments and effective capability supply. An individual with high intellectual endowments need not choose a job with high intellectual content: The choice depends on individual preferences. The implication for the specification of earnings equations is important: Earnings should not be related to an individual's available levels of capabilities, but to effectively supplied levels. Under a number of restrictions yielding a smoothly functioning perfect implicit market for capabilities, these effective supply levels will be exactly equal to the levels required in the individual's job. Hence, earnings should be related to required capability levels in the individual's job. Note that this specification diverges from the common one, where the individual's available ability levels are used. As long as there is any possibility that individuals end up in jobs that do not exactly match their ability levels (i.e., entail under-or overutilization), this specification is inadequate.[2]

The existing literature on hedonic prices, to which this study obviously relates, usually assumes indivisibilities and inseparabilities in the characteristics bundles that make up particular commodities (see Lucas, 1977; Rosen, 1974). This rules out arbitrage equating implicit prices and generally leads to unpredictability of the shape of the commodity price function. However, boldly assuming indivisibilities and inseparabilities to be irrelevant implies a commodity price equation that is linear in characteristics. This means, in the present application to the labor market, that differences in job wage rates will be interpreted as reflecting capability requirements, evaluated at given capability prices. The potential effects of indivisibilities and inseparabilities on implicit prices are thus suppressed. In itself, this would seem a very adequate first step, since the latter effects may be expected to be of secondary magnitude relative to the former, but obviously such linearity of the earnings equation is worth testing. This is done here with respect to the effects of inseparability.

The definition of capabilities stated above has no operational detail and should be given further empirical content. This is one of the questions to which this paper is addressed: Which capabilities are relevant in the labor market and which capabilities are paid for? Occupational psychologists and job analysts have studied the first question in particular (see e.g., Roe, 1956; Thorndike and Hagen, 1959). Much of their knowledge has gone into the Dictionary of Occupational Titles, which specifies required worker characteristics for thousands of jobs in the United States. This data bank was employed in research to help find the relevant capabilities. Hence, the purpose of this paper can be stated as follows: to find the required capabilities relevant in the labor market, estimate their implicit prices, and check whether inseparabilities can be ignored. In addition, since criticism of the variable with the most extensive theoretical underpinning--i.e., length of education (formal and informal)--has been mounting (see e.g., Blaug, 1976 or Cain, 1976), and in particular, questions on the content of the schooling process have come up (see Gintis, 1971; Thurow, 1975; Spence, 1974), if the capabilities structuring the labor market can be unveiled, then perhaps this also will hint at the nature of the contribution of schooling.

2. DATA

Data were collected from two different sources, the U.S. Census and the Dictionary of Occupational Titles (DOT). The first gives income data, and the second gives required worker characteristics. The match was established by comparing names of jobs: Jobs with identical titles in both data sources were assumed to be identical, so a job was included if the same title in the Census could be located in the DOT. Since the 1950 and the 1960 Censuses employed the same classification scheme, using these years allowed a comparison of changes over time.

The income concepts used in the Census do not exactly match the required concept--earnings are needed--but the Census uses a broader concept. The 1950 Census provides information on the median income in 1949 of the experienced male civilian labor force who worked 50-52 weeks of that year. Income means income before deductions, excluding income in kind and including wages and salary income, net farm income, net income from rents, boarding, etc., royalties, interests, dividends, pensions, etc., and alimony, etc. The 1960 Census gives median earnings in 1959 of the experienced male civilian labor force, aged 25-64 years, who worked 50-52 weeks of that year. Earnings, then, includes wages or salary income and self-employment income. The 1959 income concept comes closest to the desired concept of labor income, although it still includes some income from capital in the self-employment income. For this reason, attention is focused primarily on the 1959 data. Note that the inclusion of more capital and transfer income elements in the 1949 data may bias the results, thus affecting the validity of the comparison over time. This would be the case if nonlabor income were correlated with labor income. Since both transfer incomes and capital incomes are included, tending to concentrate at opposing ends of the distribution, the impact may be limited.

Required worker characteristics are taken from the DOT. The selection of the variables is in line with the definition of capabilities, i.e., productivity related characteristics. Table 1 gives the data selected and their scales of measurement. Note that DOT measures some variables "upside down," with the highest level obtaining the lowest number. For full details on these variables, the reader is referred to the DOT. Some essential information is reproduced in the Appendix.

In all, the matching procedure generated 239 observations: 40 from professional and technical workers, 49 from managers, 9 from sales workers, 16 from clerical workers, 61 from the crafts, 32 from operatives, 6 from laborers, 2 from farmers, 2 from farm laborers, and 22 from service workers. Coverage of the sample is quite good (the jobs held by 2/3 of the labor force are included).

3. INCOME DISTRIBUTION BY CAPABILITY LEVEL

The data that have been collected permit a very crude glance at income distributions by capability level as follows. For each level of a capability, a frequency distribution of median incomes can be given. Frequencies are counted as the number of workers in jobs for which the given capability level applies: For any capability level, jobs that require this level can be identified, and for these jobs both number of workers and median income level are given. The frequency distribution so obtained is a frequency distribution of median incomes.

Table 1

Measures of Job Characteristics

Symbol	Description	Value Lowest Level	Value Highest Level	Number of Intervals
DATA	degree of complexity in relation to data	8	0	9
PEOPLE	degree of complexity in relation to people	8	0	9
THINGS	degree of complexity in relation to things	8	0	9
GED	level of general education development	1	6	6
SVP	duration of specific vocational preparation	1	9	9
G	intelligence	5	1	5
V	verbal aptitude	5	1	5
N	numerical aptitude	5	1	5
S	spatial aptitude	5	1	5
P	form perception	5	1	5
Q	clerical perception	5	1	5
K	motor coordination	5	1	5
F	finger dexterity	5	1	5
M	manual dexterity	5	1	5
E	eye-hand-foot coordination	5	1	5
C	color discrimination	5	1	5

Table 2--Continued

Symbol	Description	Value Lowest Level	Value Highest Level	Number of Intervals
LEAD	leadership: situations involving the direction, control and planning of an entire activity or the activities of others	0	1	2
SALES	salesmanship: situations involving influencing people in their opinions, attitudes or judgments about ideas or things	0	1	2
HEAVY	job involves heavy or very heavy physical work	0	1	2
INDEP	independence: self-employed workers (management functions)	0	1	2

Source: U.S. Department of Labor, 1965.

Note: The variables G to M are referred to jointly as the GATB-variables (General Aptitude Test Battery).

Table 2 presents the results. The income intervals have been measured in 1959 dollars; 1949 interval boundaries were calculated by applying the ratio of the mean of the income observations in 1959 to 1949. Due to the crude nature of these measurements, interpretations should be made cautiously: This is a frequency distribution of median incomes. Some capability levels have quite low frequencies and this affects the reliability of the estimated income distribution.

The general picture that emerges from Table 2 is a positive association between level of required capability and level of income. This shows as an increase in frequencies of higher income intervals when higher levels of capability requirements are considered (the picture is somewhat clearer in a tabulation with a finer income interval), particularly in the case of G or GED.

Deviations from this general effect also occur. At higher capability levels, the dispersion sometimes tends to increase, and in particular, the frequency of the lowest income interval may rise with rising capability levels. This is illustrated by V, S, and other capabilities. In some cases, the dispersion first diminishes and then increases, as with DATA and THINGS.

Considering the changes between 1949 and 1959, it appears that in most cases there is a rather strong reduction in the frequencies of the lowest income interval and a rather weak increase in the frequencies of the highest income intervals. Frequency reductions are predominant in the \$ 3000-6000 income interval; frequency increases dominate the \$ 6000-9000 interval. Such generalizations about changes over time based on a small number of intervals are dangerous, however, since the intervals are sufficiently wide to conceal quite different patterns of movement at a finer breakdown.

Table 2

Income Distribution by Capability Level
(income in thousands of 1959 dollars)

		G = Intelligence				V = Verbal				N = Numerical				S = Spatial			
		<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9
5	1949									70.8	24.9	4.3					
	1959									64.4	35.6						
4	1949	67.5	27.0	5.5		15.0	77.8	7.2		.7	81.6	17.7		9.9	53.3	29.7	7.2
	1959	61.5	38.5			11.3	78.6	10.2		2.0	73.1	24.8		7.7	50.8	32.4	9.1
3	1949	24.0	61.3	14.6	.1	33.5	48.1	18.2	.1	27.5	46.5	24.7	1.3	41.4	36.5	21.9	.2
	1959	15.8	57.3	26.7	.2	21.9	41.2	36.6	.3	17.7	42.4	38.0	1.9	26.3	39.1	34.3	.3
2	1949		46.3	53.7			42.0	58.0			16.7	32.1	51.2		87.5	12.5	
	1959		39.7	60.3			37.7	62.3			10.6	40.6	48.8		61.9	38.1	
1	1949		8.9	54.2	36.9		10.5	46.1	43.4		2.5	95.2	2.3		15.3	58.1	26.6
	1959		8.5	53.5	38.0		9.8	46.3	43.9		2.3	88.1	9.6		2.3	66.7	31.0

		P = Form perception				Q = Clerical				K = Coordination				F = Finger dexterity			
		<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9
5	1949					21.1	75.9	3.0									
	1959					13.1	85.4	1.5									
4	1949	10.1	51.3	30.8	7.8	27.3	44.2	23.1	5.5	1.5	46.9	41.5	10.1	10.5	50.5	31.4	7.6
	1959	7.1	50.2	33.0	9.7	17.5	35.7	38.6	8.2	3.2	40.6	44.0	12.2	7.9	48.8	33.9	9.4
3	1949	30.2	50.4	19.3	.1		76.1	23.9		29.2	52.3	18.2	.2	27.6	51.3	20.9	.3
	1959	19.3	44.4	36.0	.2		60.1	39.9		17.9	49.1	31.9	1.1	16.9	44.4	37.5	1.3
2	1949		54.9	32.4	12.7		39.9	54.5	5.6		61.4	14.6	24.0		53.1	20.9	26.0
	1959		36.0	46.1	17.9		33.4	59.9	6.7		41.0	31.0	28.0		36.6	36.0	27.4
1	1949		24.1	75.9			22.6	77.4									
	1959		27.7	72.3			26.0	74.0									

Table 2--Continued.

	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9
	M = Manual dexterity				E = Eye-hand-foot coordination				C = Color discrimination			
5 1949					5.7	56.7	32.2	5.5	6.0	66.2	28.2	4.6
1959					4.3	45.0	43.5	7.2	4.3	57.2	32.6	5.9
4 1949	13.1	35.2	41.4	10.2		83.9	16.1			45.4	53.1	1.5
1959	7.8	34.2	45.3	12.8		78.5	21.5		.1	34.7	60.1	5.1
3 1949	24.2	57.3	18.3	.2	64.5	32.1	3.1	.2	84.6	7.1	3.2	5.0
1959	15.5	50.6	32.9	1.0	48.4	47.3	3.7	.6	76.0	10.4	4.0	9.5
2 1949		63.2	16.4	20.4			100.0			28.9	71.1	
1959		61.6	14.3	24.1		100.0				15.8	84.2	
1 1949											100.0	
1959											100.0	
	DATA				PEOPLE				THINGS			
8 1949	17.3	75.3	7.4		30.5	55.9	13.4	.2	1.3	42.7	44.7	11.3
1959	13.6	76.1	10.3		19.2	54.4	25.3	1.1	2.7	35.3	49.1	13.0
7 1949		73.1	26.9		1.7	98.3			60.4	32.0	7.6	
1959		50.8	49.2		14.6	85.4			51.5	48.5		
6 1949			100.0		.6	55.0	44.4					
1959		100.0			.4	51.2	48.4					
5 1949		100.0			6.5	12.1	31.4			100.0		
1959		100.0			8.2	2.3	89.5			100.0		
4 1949		100.0				100.0				95.5	4.5	
1959		100.0				100.0				96.2	3.8	
3 1949		81.4	18.6			12.7	87.3			98.8	5.5	
1959		64.5	35.5		.2	7.4	92.4			95.0	3.7	1.3
2 1949		79.5	20.0	.5		13.1	86.9			29.6	70.3	
1959		61.4	37.7	.9		45.5	54.5			21.1	78.9	

Table 2--Continued

	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9	<3	3-<6	6-<9	≥ 9
	DATA				PEOPLE				THINGS			
1 1949	38.1	17.0	35.2	9.8		.5	38.6	60.9	44.1	43.5	12.1	.3
1959	23.1	16.4	47.0	13.5		.7	40.8	58.8	29.8	33.6	34.8	1.8
0 1949		13.8	83.4	2.8		32.7	5.2	62.1		91.6	8.4	
1959		10.9	77.3	11.9		33.9	3.6	62.5		87.0	13.0	
	Specific Vocational Preparation				General Educational Development							
1 1949					81.1	13.9	5.0					
1959					75.0	25.0						
2 1949	33.0	40.2	26.8		1.9	93.0	5.1					
1959	28.5	33.8	37.6		1.1	95.4	3.4					
3 1949		97.6	2.4		1.6	91.1	7.3					
1959		97.9	2.1		4.7	84.2	11.1					
4 1949		100.0			28.3	41.8	29.8	.1				
1959		100.0			17.4	34.4	48.0	.2				
5 1949		56.5	43.5			19.1	44.0	36.9				
1959		50.5	49.5			18.9	44.6	36.6				
6 1949	54.1	19.2	26.7				95.2	4.8				
1959	40.1	20.9	39.0				81.9	18.1				
7 1949		73.4	22.6	3.9								
1959		52.0	43.3	4.7								
8 1949		6.0	63.1	31.0								
1959	.1	4.8	62.6	32.6								

Although these results are very interesting, the information they yield is still too complex to draw precise inferences about capability prices. This requires different techniques, which are applied next.

4. FACTOR ANALYSIS

The available data set contains 20 variables. Considering the definitions of these variables one may doubt whether they all reflect truly different capabilities. For example, manual and finger dexterity may be expected to be very similar, as are general intelligence and general educational development. The intercorrelation matrix underscores this view and it was decided to search for more basic underlying capabilities by applying factor analysis.

Factor analysis was applied to the GATB variables and to two larger selections: one including GED and SVP, and one covering DATA, PEOPLE, THINGS, E, and C. The motivation for including these variables requires some elaboration. GED, by definition, is a variable that comes very close to the intellectual GATB variables such as G, V, or N, and therefore cannot be taken as a really different capability. Although SVP is also used as a separate variable (in relation to the human capital view that the time dimension of it is predominant), it could equally well be argued that SVP indicates a particular capability, namely the capability of being trainable. In that case it is quite adequate to include it in the set of variables to be subjected to factor analysis. The variables DATA, PEOPLE, and THINGS have been included to investigate their role within a set of other variables: Will they merge with these variables to capabilities that have already been discovered within this set or will they bring out new relevant capabilities?[3]

The results of factor analysis on these alternative selections of variables are presented in Table 3. The outcome is very well structured with the explanatory variables divided quite clearly into two distinct factors: the intellectual capability and the manual capability. The intellectual factor is composed of high loadings of G, V, N, and Q; the manual factor is composed of high loadings of K, F, and M, and of S and P. This latter factor therefore embraces more than just physical dexterity; it also has a dimension of understanding, as reflected through S and P.

Comparison of the three alternative analyses brings out the stability of the structure: The loadings of previously included variables are barely affected and the new variables harmoniously merge into the dichotomy. GED and SVP join the intellectual factor, underscoring the interpretation of their character given above. The 3-digit code--DATA, PEOPLE, THINGS--divides between the two factors. The capability of dealing with data at increasing levels of complexity may be viewed as just another dimension of the intellectual factor; the capability of dealing with things is another manifestation of the manual factor.

The only disturbance comes from factor 3 in the third factor analysis. It is loaded highly with color discrimination, and at a lower level, the loading on PEOPLE stands out. This means that the third factor does not indicate a meaningful capability, but identifies an apparent combination of required capabilities that do not appear to stem from a meaningful latent variable. The third factor is therefore ignored in the analyses that follow.

Table 3

Three Alternative Factor Analyses
(rotated factor loadings)

Variables	Factor 1	Factor 2	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3
DATA					-.915	-.039	-.019
PEOPLE					-.526	.390	.437
THINGS					.168	-.862	-.340
GED			.930	.163	.929	.179	.003
SVP			.790	.229	.790	.318	.319
G	-.015	-.923	-.917	.040	-.906	.076	.086
V	.045	-.950	-.931	.105	-.937	.143	.104
N	-.262	-.893	-.903	-.208	-.886	-.227	-.030
S	-.834	-.232	-.310	-.823	-.302	-.862	-.007
P	-.915	-.143	-.180	-.902	-.177	-.879	.207
Q	.183	-.706	-.636	.233	-.633	.275	.103
K	-.892	.182	.150	-.899	.135	-.810	.364
F	-.936	.011	-.030	-.933	-.047	-.846	.368
M	-.883	.196	.146	-.892	.139	-.813	.378
E					.100	-.157	.375
C					-.222	-.309	.771
CUM % eigen-value	.455	.810	.440	.796	.365	.674	.764
identification	"manual"	"intell."	"intell."	"manual"	"intell."	"manual"	?

Note: The IBM-SSP computer routine was used for estimation, and factors were estimated from principal components. When rotating, only factors witl eigenvalue greater than 1 were retained, and orthogonality was maintained.

Before explaining income differences, note that factor analysis is only a formalized procedure to interpret relations between variables. It does not prove the existence or nonexistence of particular capabilities. For example, that SVP, or the variables DATA and THINGS, merge very well with factor analyses on a reduced data set that does not include these variables, does not prove that the variables do not measure something that is relevant on its own. The only conclusion from such a factor analysis is that the observations do not conflict with the hypothesis that there are two basic, latent factors which are independent of each other and which can generate the correlations between the variables actually observed.

5. EXPLAINING INCOME DIFFERENTIALS

Factor analysis has clearly brought out the relevance of two main factors for structuring the labor market. Are these factors also relevant when it comes to structuring income distribution? To this end, multiple regressions were run in various alternative combinations of explanatory variables. Results are presented in Table 4. Most adequate for investigating the question is the specification that employs 1959 income as the dependent, and that measures intellectual and manual capabilities from the GATB variables only, thus leaving room to study other variables separately.

The regressions start from a specification using the factors intellectual and manual only. The equation explains 36% of the income variance; both variables have the proper positive sign, both are significant, and the price of intellectual is about three times that of manual. Clearly, the two capabilities are relevant for income differentials.

Table 4

Regressions for 1959 Income: Factor Analysis Applied to GATB
(t-values in parentheses)

No.	Dependent	Intercept	Intellectual	Manual	GED	SVP	SVP[a]	DATA[b]	PEOPLE[b]	THINGS[b]	LEAD	SALES	HEAVY	INDEP	$\bar{R}$	F	$\bar{S}_e$
(1)	Income	5739	1070	309											.600	66.53	1489
			(11.08)	(3.18)													
(2)	Income	7322	877	208				-29	-167	-77					.620	29.16	1470
			(4.56)	(1.50)				(.46)	(2.90)	(1.48)							
(3)	Income	7340[c]	633	-89				-95	-174	-84	-273	674	373	803	.651	16.78	1438
			(2.42)	(.43)				(1.33)	(2.84)	(1.39)	(.70)	(1.60)	(1.06)	(2.08)			
(4)	Income	2196			802	112									.578	59.09	1520
					(5.53)	(1.46)											
(5)	Ln Income	8.599	.106	.058											.545	49.84	.301[d]
			(9.52)	(2.97)													
(6)	Income	5080	900	244		121									.607	45.90	1482
			(6.70)	(2.39)		(1.81)											
(7)	Income	5380	966	222			.466								.618	48.62	1466
			(7.33)	(2.22)			(2.92)										
(8)	Ln Income	8.536	.150	.043			.00008								.562	36.16	.298[c]
			(6.25)	(2.13)			(2.53)										

R=multiple correlation coefficient; F=F-test statistic; $\bar{S}_e$ = standard error of estimate (both corrected for degrees of freedom).

Notes: [a]Time-equivalent scale; see Appendix.

[b]Measured upside down; expected sign negative.

[c]Equation also contains interaction between intellectual and manual, as in note a, Table 6.

[d]The product of $\bar{S}_e$ and geometric mean of y equals 1633.

[e]The product of $\bar{S}_e$ and geometric mean of y equals 1617.

The dataset can be used to search for other capabilities that are rewarded, in particular, non-cognitive ones (the factor Intellectual is a cognitive capability; the factor Manual is a mixture of a cognitive capability as represented through P and S, and such clearly non-cognitive variables as K, F, and M). To this end, the second and the third equations are reproduced. Judging from t-ratios, these regressions bring out the relevance of PEOPLE and of INDEP. The latter variable cannot be taken as a capability without qualifications. It was measured as a dummy distinguishing employees from the self-employed. The capability of operating independently, making decisions, may be equally relevant for many salaried positions, and for proper representation the degree of independent operation should also be measured for this category. In fact, INDEP only identifies the jobs where the individuals also accept the financial consequences of their activities. Hence, INDEP may catch the return to accepting risk, or to capital.

A better case can be made for the variable PEOPLE. It represents the level of complexity in relation to other people and therefore may be interpreted as a measure of a social capability. The relevance of social capability is also brought out by the result for SALES, which borders on significance. SALES identifies job situations where other people have to be convinced. Jointly, the results on PEOPLE and on SALES are taken as evidence that a social capability is also relevant for explaining earnings differentials.

Two other aspects of the third equation merit discussion. First, LEAD comes out as an insignificant variable with the wrong sign. This result may be due to poor measurement. Counting the number of workers in jobs for which LEAD = 1, 40% of the labor force would exert leadership, but this could never be true of leadership in a restricted sense. It therefore cannot be considered a very adequate variable. Second, the results with respect to HEAVY are interesting. In some specifications (not reproduced here) it comes out with the wrong sign. However, if the factor analytic capabilities are included the sign is rightly positive (although the coefficient is insignificant). This indicates that the classical compensatory argument with respect to heavy work is valid on a ceteris paribus basis: It applies only when other factors are held constant. Omitting the c.p. clause naturally produces the wrong sign, since heavy work in the physical sense is associated with low paying jobs (compare Lucas, 1977, who obtained the same result with similar data).

Equation (4) indicates that a good deal of variance can be explained from GED and SVP. This is in line with the dominance of the intellectual capability to which both variables are strongly related. Equation (5) reports on the effect of a logarithmic transformation of income. The proper standard for comparative performance is the standard error in the linear specification relative to the standard error multiplied by the geometric mean of the dependent in the log-specification (see Sargan, 1964). Since the log-specification has a 10% larger standard error, there is no reason to prefer it to the linear model.

The last three equations of Table 4 refer to the human capital specification. Human capital theory promotes an earnings equation that explains the log of earnings from the time spent in training. Training time is not often measured independently, but the DOT-data do so. To stress the time-dimension, the scale was adjusted to reflect days of training time (the new values correspond to the midpoint of the interval; see Appendix). Note first of all that adding SVP (in the original measurement) does not add explanatory power to the equation, and barely affects estimated coefficients or t-ratios. Using the time-equivalent scale rather than SVP yields similar conclusions, although explanatory power increases somewhat. But if the log of income is used as the dependent variable instead of income itself, explanatory power diminishes (the standard error increases). Hence, in the dataset there is no gain in employing the human capital specification.[4]

To ascertain relative prices of different capabilities, standardization is called for. To this end, Table 5 presents estimates of β-coefficients for different variables and different specifications.[5] Intellectual capability comes out as the capability with the highest price; social capability, as reflected in PEOPLE and SALES, comes next; manual capability appears to carry the lowest rewards.

Next, consider interaction. Recall from Section 1 that the theory employed leads to an earnings function that is linear in regard to capabilities, a result that is intimately tied up with the assumption that capabilities are separable. The interaction term was devised to test the assumption. Interaction is measured by a dummy variable, which assumes the value, 1 if a number of variables simultaneously score beyond a certain critical level, the value 0 otherwise. The dummy thus measures the contribution to income differences of simultaneously high requirements for a number of variables. Such interaction terms have been specified for a number of combinations. Moreover, two continuously valued interaction terms were defined. The first equals the product of the manual and the

Table 5

β-coefficients for Certain Capabilities, 1959

	Factor scores derived from		
Variable	GATB[a]	GATB, GED, SVP[b]	GATB, E, C, GED, SVP, DATA, PEOPLE, THINGS[c]
Intellectual	.34	.41	.57
Manual	.05	.06	.05
THINGS	.14	.13	
PEOPLE	.21	.18	
SALES	.09	.12	.10

Notes: [a]Regression equation includes DATA, LEAD, HEAVY, INDEP, and interaction as in Table 6, note a.

[b]Regression equation as in note a, plus some insignificant other interaction dummies.

[c]Regression equation as in note a (except DATA), plus another insignificant interaction dummy.

intellectual factors and thus allows for substitution (medium scores on both factors are valued equal to a high score on one and a low score on the other). Substitution is prohibited in the other continuous interaction variable, the minimum of the two factor scores. In this case, interaction is very strong: The importance of the factors cannot surpass the value of the lowest reqiurement. In other words, a high requirement of one factor in only relevant (with respect to earnings) if the other factor is also required at a high level.[6]

Table 6 summarizes the evidence on the relevance of interaction: Most species are insignificant and/or have the wrong sign (they were all defined such as to require a non-negative coefficient). Only one interaction comes out significantly: the interaction between Manual and Intellectual, provided it is defined with a sufficiently high critical level. This interaction identifies 16 professional occupations.

Do these results reject the assumption of separability? Although it casts some doubt, the argument is not unequivocal. First, note that explanatory power is not enhanced by the interaction term. Comparing the first equation in Table 6 with the first in Table 4, $\bar{R}^2$ rises only from .36 to .37, and the standard error of estimate is reduced by only .5%. Next, the result may reflect a temporary situation rather than a permanent one. In the counterpart of the first equation, 1949 income, interaction was not significant.[7] Clearly, the question requires further empirical research.

Since data were collected for 1949 as well as 1959, changes in capability prices over time may be studied. Capability requirements were only observed once, in the 1965 edition of the DOT. Hence, implicitly, capability requirements are assumed constant for given jobs. Obviously,

Table 6

Testing for Interaction, 1959

Factor Analysis Applied to	Intell Man[a]	Intell Man[b]	Intell Man[c]	Min. Intell Man[d]	Intell Man Lead[e]	Intell Lead[f]	Intell Data[g]	Intell People[h]	Other Variables Included in the Regression
	1)	2)	3)	4)	5)	6)	7)	8)	
GATB	991 (1.90)								intell, man
		-66 (.13)							idem
	1915 (3.24)								intell, man, GED, SVP, DOT 3, LEAD, SALES, HEAVY, INDEP.
	1967 (4.57)								intell, SVP, PEOPLE, SALES, INDEP.
GATB, GED, SVP	1220 (2.50)								intell, man
			10 (.10)						idem
				-1194 (3.60)					idem
						582 (1.57)			idem
	1730 (3.08)								intell, man, DOT 3, LEAD, SALES, HEAVY, INDEP
			83 (.72)						idem

Table 6--Continued.

Factor Analysis Applied to	Intell Man[a]	Intell Man[b]	Intell Man[c]	Min. Intell Man[d]	Intell Man Lead[e]	Intell Lead[f]	Intell Data[g]	Intell People[h]	Other Variables Included in the Regression
				-1196 (3.27)					idem
						621 (1.39)			idem
	2488 (2.97)						-1014 (1.21)	-417 (.55)	Idem
GATB, E, C, GED SVP, DOT 3[i]	1905 (4.54)								intell, SALES, HEAVY, INDEP.
	1923 (3.11)				-470 (.64)				intell, man
	2156 (3.34)				-262 (.30)				intell, man, LEAD, SALES, HEAVY, INDEP.

All variables were measured such as to have a non-negative expected sign.

Notes: [a]Intellectual and manual each at least 1 standard deviation above the mean.

[b]Intellectual and manual each at least .5 standard deviation above the mean.

[c]Intellectual X manual.

[d]Minimum of intellectual and manual.

[e]Intellectual and manual as in note a, and leadership = 1.

[f]Intellectual as in note a, and leadership = 1.

[g]Intellectual as in note a, and data = 0 (highest level).

[h]Intellectual as in note a, and people = 0 (highest level).

[i]DOT 3 = DATA, PEOPLE, THINGS.

exact matching in time of income and capability prices would have been preferable, but no such data are available.

Table 7 collects evidence on changes in relative prices as measured through different specifications. All tend to tell the same tale. The price of the intellectual capability is lagging behind, the social capability (PEOPLE) has increased somewhat in relative price, and, in particular, the manual capability has become worth more. Although these results seem plausible, it is still important to recall the differences in income concept; it cannot be entirely ruled out that these changes reflect a bias.

6. CONCLUDING REMARKS

This paper has stressed the need to relate earnings to variables describing an individual's job rather than the individual. In searching for capabilities structuring the labor market, two capabilities emerged from a factor analytic approach: intellectual and manual capability. As for earnings, a third capability also proved relevant, i.e., a social capability. The intellectual capability carries by far the highest rewards, followed by social, with the manual trailing behind. There is some evidence that the relative price of manual capability increased between 1949 and 1959. Inseparability of capabilities, producing an interaction effect in the earnings equation, could not be entirely eliminated, and this issue needs further testing.

In concluding, some remarks should be made to relate the present work to Thurow and Lucas (1972). They matched the DOT data with income

25

Table 7

Changes in Relative Prices: 1959 Regression Coefficients
Divided by 1949 Coefficients

Factor Analysis on	Intercept	Intellectual	Manual	DATA	PEOPLE	THINGS
GATB	1.62	1.67	2.92			
	1.63	1.81	5.20	.76	1.86	1.75
GATB, GED, SVP	1.62	2.67	3.57			
	1.62	1.81	3.68	.79	1.87	1.50
GATB, GED, SVP DATA, PEOPLE THINGS, E, C	1.62	1.65	3.96			

observations from the Survey of Economic Opportunity by applying a matrix cross-classifying the occupations in the two datasets. Regressions of hourly earnings were run on the DOT variables (more variables than used here) and on personal characteristics for large samples (e.g., 8700 observations for white males). They obtained negative prices for a number of characteristics, and their study differs from this one in a number of other aspects. First, they did not structure the variables (such as applied here with factor analysis), but used them all simultaneously in a multiple regression, and therefore there may be a serious problem of multicollinearity. In this regard it is relevant to mention their results with respect to some of the subsets of the variables as determined here on the basis of intercorrelation (and factor analysis). Intelligence (G), obtains a negative coefficient in most of their regressions, but numerical ability (N) a positive; also, in the one case that G has a positive coefficient, N has a negative one. This points to the interdependence of the estimates and may serve as a warning for multicollinearity. A similar effect occurs with the variables K, F, and M, which in the sample analyzed here appeared highly correlated and factor analytically related to the same latent factor. In Thurow and Lucas's sample, negative coefficients occur on the variables K, F, and M, but never on all three variables simultaneously.

Second, as far as the GATB aptitudes are concerned, measurement did not specify the required level as given in the DOT, but the probability of the job requiring that level of the ability possessed by the top 20% of the population. This means that only earnings associated with top level requirements were estimated, instead of assuming a continous relation between requirement and reward. In all, the difference in method and underlying model makes it hard to compare the results of the two studies.

APPENDIX A

The DOT-data

1. DATA, PEOPLE, THINGS

The DOT classifies a job according to its place in three hierarchies, arranged "from the relatively simple to the complex in such a manner that each successive relationship includes those that are simpler and excludes the more complex."

2. GED: General Education Development

GED embraces those aspects of education (formal and informal) that contribute to the worker's a) reasoning development and ability to follow instructions, and b) acquisition of "tool" knowledges, such as language and mathematical skills. It is education of a general nature, which does not have a recognized, fairly specific, occupational objective.

3. SVP: Specific Vocational Preparation

SVP indicates the amount of time required to learn the techniques, acquire information, and develop the facility needed for average performance in a specific job-worker situation. The scale used and its transformation to SVP (Table 4) are as follows.

SVP-value	Meaning	SVP*-value
1	Short demonstration only	1
2	< 30 days	15
3	30 days - < 3 months	60
4	3 months - < 6 months	135
5	6 months - < 1 year	270
6	1 year - < 2 years	540
7	2 years - < 4 years	1080
8	4 years - < 10 years	2520
9	> 10 years	4320

4. GATB, E, C: General Aptitude Test Battery, Eye-hand-foot coordination, Color discrimination

GATB, E, and C are specific capabilities and abilities required of an individual in order to learn or perform adequately a task or job duty. The amount required is expressed in terms of equivalent amounts possessed by segments of the general working population: 1 = top 10%, 2 = next 23 1/3%, 3 = middle third, 4 = next 23 1/3%, 5 = lowest 10%.

NOTES

[1]An exception is Lydall (1968).

[2]If there is no smoothly functioning market for capabilities (e.g., if indivisibilities are significant), the earnings function needs modification and may contain the available levels of capabilities as well; this is treated in Berkouwer, Hartog and Tinbergen (1978). Note that Thurow and Lucas (1972) also stress the importance of job characteristics rather than the individual's characteristics, but for different reasons.

[3]The variables E and C were also added as input into the factor analysis; *a priori*, they are not expected to be very relevant for either the labor market structure or for earnings differentials.

[4]The same conclusion is reached by Ten Cate (1977), using comparable Dutch data.

[5]A β-coefficient is a regression coefficient obtained when all variables (dependent and independent) have been standardized to unit variance.

[6]Note that these remarks apply to the interaction between the factors; all regression equations include the factor requirements separately as well.

[7]In some other specifications it was significant.

REFERENCES

Berkouwer, J., Hartog, J., and Tinbergen, J. 1978. Alternative specifications of earnings equations (Discussion Paper 7804/G). Institute for Economic Research, Erasmus University, Rotterdam.

Blaug, M. 1976. The empirical status of human capital theory: A slightly jaundiced survey. Journal of Economic Literature, 14 (3), 827-855.

Cain, G. C. 1976. The challenge of segmented labor market theories to orthodox theory. Journal of Economic Literature, 14 (4), 1215-1257.

Gintis, H. 1971. Education, technology and the characteristics of worker productivity. American Economic Review, Papers and Proceedings, 61 (2), 266-279.

Hartog, J. 1978a. Personal income distribution: A multicapability theory. Ph.D. thesis, Erasmus University, Rotterdam.

Hartog, J. 1978b. On the multicapability theory of income distribution. European Economic Review, 10 (2), 157-171.

Lucas, R. E. B. 1977. Hedonic wage equations and psychic wages in the returns to schooling. American Economic Review, 67 (4), 549-558.

Lydall, H. 1968. The structure of earnings. Oxford: Clarendon Press.

Roe, A. 1956. The psychology of occupations. New York: Wiley.

Rosen, S. 1974. Hedonic prices and implicit markets: Product differentiation in pure competition. Journal of Political Economy, 82 (1), 34-55.

Roy, A. D. 1951. Some thoughts on the distribution of earnings. Oxford Economic Papers, 3, 135-146.

Sargan, J. D. 1964. Wages and prices in the United Kingdom: A study in econometric methodology. In P. E. Hart, G. Mills and J. K. Whitaker (Eds.), Econometric Analysis for National Economic Planning. London: Butterworths Scientific Publications.

Spence, A. M. 1974. Market signaling. Cambridge: Harvard University Press.

Ten Cate, A. 1977. Beroepen: vereisten en beloning (Discussion Paper 7707/G). Rotterdam: Institute for Economic Research, Erasmus University.

Thorndike, R. L., and Hagen, E. 1959. Ten thousand careers. New York: Wiley.

Thurow, L. 1975. Generating inequality. New York: Basic Books.

Thurow, L., and Lucas, R. E. B. 1972. The American distribution of income: A structural problem (Joint Economic Committee, Congress of the United States). Washington, D. C.: U.S. Government Printing Office.

Tinbergen, J. 1956. On the theory of income distribution. Welt-wirtschaffliches Archiv, 77, 156-175.

U.S. Bureau of the Census. 1950. Occupational characteristics, Special reports, Census of population. Washington, D. C.: U.S. Government Printing Office.

U.S. Bureau of the Census. 1960. Occupational Characteristics, Subject reports, Census of population. (Final Report PC (2)-7A). Washington, D. C.: U.S. Government Printing Office.

U.S. Department of Labor. 1965. Dictionary of occupational titles (Vols. 1 and 2) (3rd ed.). Washington, D. C.: U.S. Government Printing Office.