

A complete list of Working Papers
on the last pages

No. 29, 1980

**A PROCEDURE FOR TESTING
THE SIGNALLING HYPOTHESIS**

by

James W. Albrecht

July, 1980

This is a preliminary paper. It is intended for private circulation, and should not be quoted or referred to in publications without permission of the author. Comments are welcome.

Abstract

This paper develops a procedure for testing the signalling hypothesis as advanced by Spence, et. al. The approach used is to examine directly the question of whether employers use education for purely informational purposes in their hiring decisions. An application of the method to a recruitment by the Swedish auto manufacturer Volvo is presented.

James W. Albrecht
1018 International Affairs
Dept. of Economics
Columbia University
New York, N. Y. 10027
USA

A Procedure for Testing the Signalling Hypothesis*

1. Introduction

The signalling model of the returns to education as developed by Spence (1974), Arrow (1973) and Stiglitz (1975) represents an important theoretical contribution to the economics of information, but whether this contribution is of significant empirical consequence is an open question. This paper develops and applies a general method for addressing this question.

The signalling interpretation of the returns to education depends upon employers' lack of information about job applicants. Workers (applicants) are assumed to have a good idea about their productivities, but, a priori, employers are not. If the less productive cannot be induced to admit to that fact, then the employer considering job applicants will be forced to "estimate" applicants' productivities.

It is suggested that educational background may serve as an ideal observable trait for the employer to use to infer other, unobservable traits related to productivity. That education can be so used depends

*This paper has gone through several versions and two data sets. The earliest version was presented at the 1974 Econometric Society meetings in San Francisco. The guidance and encouragement of Roy Radner on the early versions is gratefully acknowledged. The latest version was presented at a conference on Labor Market Issues in Sweden held at the Industrial Institute for Economic and Social Research, Stockholm in July 1979. The comments of several of the conference participants, especially Bertil Holmlund, have been incorporated into the text.

upon the assumption that the cost of education varies inversely with productivity. Under this assumption only the inherently more productive will find extra education worthwhile. Employers' initial beliefs that the educated are more productive will be self-fulfilling.

This basic objection to the signalling hypothesis is that the educational screen is a costly one. Ought not there exist less expensive alternative mechanisms to elicit information about productivities from applicants?

It is sometimes asserted, for example, that any signalling component to the rewards to education would be eroded by the establishment of "testing firms." Alternatively, firms may be able to structure their promotion policies in such a way as to deter applicants from misstating their qualifications; that is, applicants may be induced to self-select into the proper job slots. These arguments, however, lack any empirical basis.

My approach to the signalling hypothesis will be to examine directly the question of whether employers reward education for purely informational purposes in the hiring decision. The role of education in the hiring decision will be decomposed into a pure "productivity component" and a pure "information component." This is most naturally done within a 2-way analysis of covariance framework with interactions between education and "information."

In the next section I develop the statistical procedure for testing the signalling hypothesis. Then, in the third section, I present an application

of this method to a recruitment by the Swedish auto manufacturer Volvo. The results of this application provide both an illustration of the procedure and some substantive evidence about the signalling hypothesis. Finally, in a concluding section, I summarize the method and relate my procedure to another approach presented in Riley (1979).

2. A General Procedure

Suppose an employer is considering applicants for a position who can be characterized by their educational background and by their "information level," i.e., the amount of a priori information the employer has about them. According to the signalling hypothesis, employers need to use education as a source of information about applicant productivities, i.e., applicants cannot be induced to properly self-select by some cheaper means. Therefore, if the signalling hypothesis is valid, employers will be forced to rely more heavily on education when considering those applicants about whom they have the least information. The test procedure presented below is an exploitation of this simple idea,

Typically "information level" will be a qualitative variable, and often educational attainment will be as well. Let $i = 1, \dots, I$ index educational categories, and let $j = 1, \dots, J$ index informational categories. The k^{th} individual in the $(i, j)^{\text{th}}$ cell has observable characteristics X_{ijk} . Assume that the (lifetime, discounted, etc.) marginal product ($= Z_{ijk}$) of this applicant as perceived by the prospective employer can be expressed as a linear combination of these characteristics plus a $N(0, \sigma^2)$ error term. That is,

$$Z_{ijk} = X_{ijk}\eta + U_{ijk}, \quad \text{where } U_{ijk} \text{ is } N(0, \sigma^2) \quad (1)$$

The employer's decision problem can be modelled as one of accepting only those applicants whose perceived marginal product exceeds a critical value w . Then, the probability that the k^{th} applicant in the $(i, j)^{\text{th}}$ cell will be accepted can be written as

$$\begin{aligned}
P_{ijk} &= \Pr(Z_{ijk} > w) = \Pr(X_{ijk}\eta + U_{ijk} > w) = \\
&= \Pr(U_{ijk} < X_{ijk}\eta - w) = \frac{X_{ijk}\eta - w}{\sigma} \int_{-\infty}^{\sigma} (2\pi)^{-1/2} e^{-z^2/2} dz \\
&= \Phi(X_{ijk}\eta^*). \tag{2}
\end{aligned}$$

where η^* is the standardized parameter vector and $\Phi(\cdot)$ is the distribution function of the standardized normal random variable.

To pursue the 2-way analysis of covariance approach, assume

$$X_{ijk}\eta^* = \mu + \alpha_i + \beta_j + \lambda_{ij} + \sum_{h=1}^H \delta_h y_{hk}, \tag{3}$$

where

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_i \lambda_{ij} = \sum_j \lambda_{ij} = 0.$$

The interpretation of the parameters is as follows:

μ = mean (standardized, perceived) productivity

α_i = main effect on productivity of being in educational category i

β_j = main effect on productivity of being in informational category j

λ_{ij} = interaction effect on productivity of being jointly in educational category i and informational category j

δ_h = effect of the h^{th} concomitant variable on productivity.

The main effects of education are the effects of educational categories averaged across all informational categories, and likewise for the main

effects of information. The interaction effect in the $(i,j)^{\text{th}}$ cell is the effect of the i^{th} level of education on the employer's perception of applicant productivity specific to the j^{th} informational category; that is, it is the effect of the i^{th} level of education above and beyond the main effect, α_i .¹

It is the interaction effects which are of principal interest. To see this it is useful to consider a simple "2x2" example. Imagine an applicant pool differentiated according to high versus low education level and high versus low information level. If employers are forced to use education for information, then the interaction effects can be expected to take on the sign pattern indicated below. We expect education

		<u>Information</u>	
		High	Low
<u>Education</u>	High	-	+
	Low	+	-

to receive a positive overall weight in the employer's assessment procedure. If part of this positive overall weight can be ascribed to an informational component, then the positive effect of education ought to be decreased in the presence of alternative information; i.e., we expect the interaction effect for high education together with high information to be negative. Analogous arguments can be made to sign the other interaction terms, but these are redundant since there is only one independent interaction parameter in this 2x2 case. Alternatively, if the employer is not forced

¹The concepts of main effects and interaction effects in 2-way analysis of variance models are lucidly discussed in Scheffé (1959).

to use education as a source of information, then the effect of education should be constant across all information levels. Thus, a test of the hypothesis that the employer does not use education for informational purposes may be expressed as

$$H: \lambda_{ij} = 0; \quad i = 1, \dots, I, \quad j = 1, \dots, J.$$

It is to be emphasized that the hypothesis of zero interaction effects is not the hypothesis that the employer is indifferent about the educational attainment of applicants, nor is it the hypothesis that the employer is indifferent about the amount of a priori information available about prospective employees. These hypotheses instead translate into hypotheses about the main effects.

Nor does the hypothesis of zero interaction effects imply that an employer's preference for applicants about whom more information is available need solely reflect a preference for more information. There may be differences in average productivity across information classes, but these differences ought to be reflected in the main effects of information, rather than in differential rewards to education. However, one must be on guard for other mechanisms that might introduce an interaction between education and information, and such alternative mechanisms are easier to imagine when information is not "neutral." The point, of course, is that one must be careful in specifying "information classes."

3. An Application

The data used in this application come from records of applicants for entry-level blue collar positions at Volvo's Torslanda auto works for the month of June 1978.² Excluding those applying for part-time work, a total of 515 applicants were considered and of these 291, or 56.5%, were hired. Data on the educational attainment and on the recruitment source of each applicant are available from these records. Educational attainment is a dichotomous variable with "high education" identified with attainment of less than the gymnasial level. The gymnasium is normally attended for 3 years in Sweden between the ages of 16-19 and roughly corresponds to the last years of senior high school plus parts of junior college in the U.S. Today the completion rate in the gymnasium is quite high, but this is a very recent phenomenon, and in this sample 42% of the applicants have not completed the gymnasium.

The information class of the applicant is identified with the source of his or her recruitment. The first recruitment source - and this is the source to be identified with greater prior information - is recommendation by a current Volvo employee; that is, the applicant has given the name of a Volvo employee who has informed him of the job opening and from whom the personnel department can solicit an evaluation. Of course, such an evaluation may not be unbiased, but it seems reasonable that the company

²These data were kindly made available to me by Göte Bernhardsson and Anne-Marie Qvarfort of the Employment Research Group in the Swedish Labor Market Department. Their report on Volvo's recruitment practices is available in mimeo as "Personalrekryteringen till Volvo-Torslandaverken, Juni 1978", Sysselsättningsutredningen, October 1978.

can take the caliber of the reference into account. The other two recruitment sources are identified with less prior information. The first of these relatively low information sources is the Swedish Labor Market Board (AF). This refers to job seekers who have searched AF's position announcements and have then come to Volvo with a notification from that Board. No active placement on AF's part is implied. Second, there are those who have simply applied in response to newspaper advertisements (plus a small group from "miscellaneous" sources). In principle, those who come via AF and those who come via advertisement are in an equally low information category. However, there is the possibility of more active placement on the part of the AF for some candidates. This potentially has both the implication of more information and the implication of a decrease in the probability of hire for those candidates since AF is more likely to make an active effort on behalf of those who are "difficult to employ." These two low information categories have been combined in the empirical results presented below.³

Besides the information about education and recruitment source, data are available on the age, the nationality, the residence and the sex of each applicant. These data are presented in Table 1. Ignoring any covariation between these variables for the moment, Table 1 indicates a preference for

³In fact, Bernhardsson and Qvarfort conjecture that some applicants recorded as recruited via advertisement may also have searched the AF position announcements. There are 2 bases for this suspicion: (i) some applicants may feel that any identification with AF hurts their chances and (ii) the fraction of applicants coming from AF seems "abnormally low."

(1) more highly educated applicants, (2) applicants in the high information category, (3) younger applicants, (4) Swedish and Finnish nationals, (5) non-Gothenburg residents and (6) males. The only surprise in the data is the preference given to those living out of the greater Gothenburg region where the plant is located. However, the relatively low number of non-Gothenburg residents (and the even lower number of females) among the applicants should be noted.

The model that has been estimated inverts equation (2) to express $\Phi^{-1}(p)$ as a constant plus a sum of main effects for education, information, age, nationality, residence and sex plus an education-information interaction. The parameters have been estimated using maximum likelihood (probit), and test statistics for assessing the significance of the main and interaction effects have been computed as -2 times the logarithm of the appropriate likelihood ratio. The test statistics are asymptotically χ^2 with degrees of freedom equal to the number of independent restrictions implied by the null hypothesis. These parameter estimates and test statistics are presented in Table 2.

The parameter estimates may be interpreted with the aid of a simple example. An applicant who (1) has a low level of education, (2) falls in the high information category, (3) is between the ages of 21-27, (4) is Swedish, (5) is a Gothenburg resident and (6) is male would be hired with an estimated probability of $\Phi(0.305) = 0.620$. An applicant with a high level of education but otherwise identical attributes would be hired with an estimated probability of $\Phi(0.569) = 0.715$ with

the change ascribable to the increase via the main effect of education (from -0.191 to +0.191) and to the decrease via the education-information interaction (from +0.059 to -0.059).

The pattern of main effects in Table 2 is in basic accord with that suggested by the raw data in Table 1. Completion of the gymnasium, Swedish or Finnish nationality and being male strongly increase the chance of getting hired, and these main effects are significant at the 1% level. Having a Volvo employee to use as a reference also increases the hire probability, but not as strongly; and the factors of age and residence, while retaining the same pattern as in the raw data, become much less important. In fact, the anomalous apparent preference for non-Gothenburg residents essentially becomes zero when the covariation between residence and other variables is taken into account. The significance probabilities for the main effects of information, age and residence (0.15, 0.35 and 0.65, respectively) are above conventionally accepted levels.

The interaction effects take on the sign pattern suggested by the signalling hypothesis, i.e., the positive effects of extra education are decreased in the presence of extra information, but these effects are quite small in magnitude. The significance probability for the education-information is only slightly less than 0.5. The hypothesis of zero interaction effects clearly cannot be rejected; that is, the hypothesis that there is no purely informational component to the preference exhibited for the more educated applicants cannot be rejected. Volvo's hiring behavior gives no support to the signalling hypothesis in this instance.

Finally, it should be noted that the results are insensitive to re-parameterization of the basic model. Alternative models have been estimated with (1) age as a continuous variable, (2) 3 information categories instead of 2, (3) interactions between education and nationality and information and nationality and (4) residence and sex suppressed as separate variables. In addition the model with residence and sex suppressed has been re-estimated by the alternative technique of "minimum normit chi-square," i.e., weighted least squares based on the cell relative frequencies, as developed by Berkson (1955). The basic conclusions remain the same.

4. Discussion

This paper has presented a procedure for testing the signalling hypothesis based on a decomposition of the role of education in the hiring decision into a pure "productivity" component and a pure "information" component. The procedure was applied to a recruitment of auto workers by Volvo, and in this instance Volvo's hiring behavior indicates no support for the signalling hypothesis. Volvo prefers applicants with more education and (weakly) prefers applicants about whom more information is available, but in the absence of that extra information no significantly different premium is attached to extra education. That is, Volvo does not appear to rely on education for purely informational purposes in the hiring process.

Of course, this same procedure could be applied to different sets of data, and one aim of this paper is to motivate the collection of

richer data sets for replication. As explained above, and as illustrated in the Volvo application, the trick is to define the concept of "information level" in a suitable way.

The procedure developed in this paper is very "micro" in the sense that it focusses on the significance of signalling at the level of the individual job and at the level of the individual employer. More "macro" approaches are also possible, and such approaches can be considered complementary to the method advocated here. In my opinion, the best of these macro approaches is presented in Riley (1979).⁴ Riley's method is based on an idea similar to that of information levels. However, instead of differentiating among applicants for a particular job according to the amount of available prior information, he divides occupations into those for which productivity may be easily ascertained versus those for which signalling might conceivably be important. A test of the signalling hypothesis is then based on a comparison of lifetime earnings functions at each level of education for those in the "screened" sector versus those in the "unscreened" sector. Using this test, Riley concludes that signalling is a significant phenomenon.

However, as Riley points out, there is no obvious best method for classifying occupations as screened or unscreened. In fact, he is forced to use ex post data analysis to perform the classification. Nor is there any

⁴ Some other empirical papers on signalling are Layard and Psacharopoulos (1974), Taubman and Wales (1973), and Wolpin (1977). Riley gives a good discussion and critique of these papers.

way to ensure the differences in earning profiles between the screened and unscreened sectors for a given education level can be solely ascribed to the screening function of education. But these practical problems are analogous to the ones which make the application of this paper's procedure difficult; namely, suitably defining what one means by "information" and ensuring that specious interactions between education and information are controlled.

To summarize, empirical analysis of the significance of signalling appears to have reached the point where well-founded techniques are becoming available. However, the data requirements imposed by these techniques have proven to be rather stringent. One advantage of the procedure and application presented in this paper is that these data requirements have been clarified, and one can hope that further applications will be possible.

Table 1 The basic data

		Applicants	Hired	Relative frequency
	Total	515	291	.565
Education:	Low	215	101	.470
	High	300	190	.633
Information:	Rec	180	110	.611
	AF	115	58	.504
	Ad	220	123	.559
Age:	≤ 20	202	129	.639
	21-27	186	104	.559
	≥ 28	127	58	.457
Nationality:	Swedish	298	182	.611
	Finnish	122	76	.623
	Other	95	33	.347
Residence:	Gothenburg	415	228	.549
	Other	100	63	.630
Sex:	Male	455	270	.593
	Female	60	21	.350

Source: Unpublished data from the Employment Research Group of the Swedish Labor Market Department.

Table 2 Probit Estimates and Significance Tests

	Maximum Likelihood Estimate	Test Statistic (χ^2)		Maximum Likelihood Estimate	Test Statistic (χ^2)
Mean	-0.145	---	Nationality		14.78*
<u>Main Effects</u>			Swedish	0.185	
			Finish	0.240	
			Other	-0.425	
Education		9.48*	Residence		0.34
High	0.191		Gothenburg	-0.016	
Low	-0.191		Other	0.016	
Information		2.18	Sex		10.98*
High	0.115		Male	0.320	
Low	-0.115		Female	-0.320	
Age		1.98			
>20	0.095				
21-27	-0.022				
>28	-0.073				
<u>Interaction Effect</u>		0.54			
HighEdxHighInf	-0.059				
HighEdxLowInf	0.059				
LowEdxHighInf	0.059				
LowEdxLowInf	-0.059				

*Significant at 1% level

References

1. Arrow, K., 1973, Higher Education as a Filter, *Journal of Public Economics*, 2, 193-216.
2. Berkson, J., 1955, Estimate of the Integrated Normal Curve by Minimum Normit Chi-Square with Particular Reference to Bio-Assay, *Journal of the American Statistical Association*, 50, 529-549.
3. Layard, R. and G. Psacharopoulos, 1974, The Screening Hypothesis and the Returns to Education, *Journal of Political Economy*, 82, 985-998.
4. Riley, J., 1979, Testing the Educational Screening Hypothesis, *Journal of Political Economy*, 87, S227-S252.
5. Scheffé, H., 1959, *The Analysis of Variance* (John Wiley, New York).
6. Spence, A. M., 1974, *Market Signalling: Informational Transfer in Hiring and Related Screening Processes* (Harvard, Cambridge).
7. Stiglitz, J., 1975, The Theory of "Screening", Education and the Distribution of Income, *American Economic Review*, 65, 283-300.
8. Taubman, R. and T. Wales, 1973, Higher Education, Mental Ability, and Screening, *Journal of Political Economy*, 81, 28-55.
9. Wolpin, K., 1977, Education and Screening, *American Economic Review*, 67, 949-958.