DISCUSSION PAPER NO. 452-R

AN EMPIRICAL INVESTIGATION OF THE RETURNS TO JOB SEARCH

Robert S. Chirinko*

October 1981

Forthcoming in The American Economic Review

* Visiting Assistant Professor, Northwestern University.
I wish to acknowledge the extensive comments and assistance offered by Stephen R. Cosslett, Walter D. Fisher, Daniel S. Hamermesh, Stephen R. King, Shelly J. Lundberg, Wesley Mellow, Dale T. Mortensen, Edward C. Prescott, George W. Wilson, two anonymous referees, and participants in the Macro/Labor Workshop at Northwestern University. All errors and omissions remain the sole responsibility of the author.

AN EMPIRICAL INVESTIGATION OF THE RETURNS TO JOB SEARCH

The assumption of diminishing returns to job search plays a crucial role in the results derived from a search model presented recently in this *Review* by John Seater (1979). This assumption is justified by noting that "spatial aspects of the search process cause diminishing returns" since the travel time required to search firms in a circular area increases faster than the area to be searched.

Besides being an important property of a class of theoretical models, returns to search may serve to explain partly the paradox of an approximately 40 hour workweek contrasted with a seven hour "searchweek" (Barron and Mellow (1979), p. 396). The interaction of diminishing returns to and positive costs of search may result in an optimal allocation of search time of only a few hours per week.

In this paper, we examine whether the assumption of diminishing returns to job search can be supported by the data. Following Seater, the returns to search are measured by the number of firms contacted, and we focus on the ability of unemployed individuals to translate search time into employee contacts. Returns to search are evaluated by noting whether the "production process" exhibits diminishing returns with respect to the input variable.

I. ECONOMETRIC MODEL

Estimating the contacts-search time "technology" by Ordinary Least Squares (OLS) is inappropriate since employee contacts, the dependent variable is integer-valued and non-negative. The possibility that the variables in the regression may be zero for some individuals precludes a logarithmic transformation of the model as a solution to the non-negativity problem.

These problems can be treated adequately in the following two-step econometric procedure. First, the number of firms contacted in a fixed period of time is assumed to "arrive" at a rate described by a Poisson distribution,

$$\mathrm{PROB}(N_i = K \mid \lambda_i) = e^{-\lambda_i} \lambda_i^K / K! \qquad (1)$$

N_i = The number of firms contacted in a given period by i
K = 0,1,2,...
λ_i = Expected number of firms that will be contacted by i (i.e., $\lambda_i = E(N_i)$, $E(\cdot)$ is the expected value operator)
i = 1,I (individuals).

In turn, the Poisson parameter for individual i, λ_i, is related generally to search time, and a set of demographic and labor market variables entered both separately and interactively with search time. In order to allow for the possibility of non-constant returns to search, a term squared in search time is included as an explanatory variable:

$$\lambda_i = E(N_i) = b_0 + b_1 * ST_i + b_2 * ST_i^2 + \sum_{j=1}^{J} b_{j+2} * Z_{i,j} + \sum_{j=1}^{J} b_{j+J+2} * Z_{i,j} * ST_i \qquad (2)$$

ST_i = The total hours spent looking for work in a given period by i
$Z_{i,j}$ = The jth demographic or labor market variable
$b_0, b_1, \ldots, b_{2J+2}$ = Parameters to be estimated.

Equation (2) generalizes Seater's equation (18), which describes a representative individual searching in a geographical plane, by recognizing that individual characteristics may affect the intercept or slope of the contacts-search time production function.

The econometric model, (1) and (2), can be estimated by the method of maximum likelihood.[1/] Gilbert (1979) has demonstrated that the maximum likelihood estimator for a Poisson process is equivalent to an Iterated Weighted Least Squares estimator (IWLS), which can be calculated easily with conventional software packages,

$$\hat{B} = (X'W^{-1}X)^{-1} X'W^{-1}N \tag{3}$$

$$V(\hat{B}) = (X'W^{-1}X)^{-1} \tag{4}$$

$B = (b_0, b_1, \ldots, b_{2J+2})$ - Kx1 vector of parameters
X = IxK matrix of explanatory variables
N = Ix1 vector of dependent variables
W = IxI diagonal weighting matrix whose non-zero elements are given by $\hat{N} = X\hat{B}$
$V(\cdot)$ = KxK asymptotic covariance matrix.

The estimator (3) is computed iteratively. First, W is set equal to the identity matrix and a standard OLS regression is performed. These initial estimates of B are inserted into W and (3) is reestimated. The process of recomputiong W and reestimating $\hat{B}$ continues until the changes in elements of $\hat{B}$ are less than a predetermined convergence criteria.[2/]

The returns to search hypothesis is stated in terms of the coefficient on ST_i^2,

$$\begin{aligned} H_0&: \ b_2 = 0.0 \\ H_a&: \ b_2 < 0.0. \end{aligned} \tag{5}$$

Given $\hat{B}$, $V(\hat{B})$, and a large sample, (5) can be evaluated using a normal distribution to determine if the null hypothesis can be rejected in favor of the alternative.

II. EMPIRICAL RESULTS

The model is estimated using two samples drawn from a special survey of unemployed individuals in the Current Population Survey (CPS), May 1976.[3/] The respondents were asked, among other questions, the number of hours spent looking for work, the method used most frequently, and the number of firms contacted during the previous four weeks. Those who classified themselves as being on layoff and not looking for another job, or unemployed and not spending any time searching have been excluded. Since participants in the special survey were also included in the CPS, we are able to assess the influence on the contacts-search time technology of a broad range of demographic and labor market characteristics.

A. Direct Search - Sample A

This sample is restricted to those respondents whose most frequent method of search was to "apply directly to employers without suggestions or referrals by anyone," and it is this type of activity that closely approximates Seater's notion of spatial search. The regression is specified with the number of firms contacted by an individual dependent on search time (ST_i), the square of search time, a constant, a dummy variable, $CITY_i$ (1 if living in a SMSA), and interactive terms, defined as search time multiplied separately by a series of variables: AGE_i (chronological age), $RACE_i$ (dummy variable, 1 if nonwhite), SEX_i (dummy variable, 1 if female), UC_i (dummy variable, 1 if receiving unemployment insurance benefits), and TL_i (dummy variable, 1 if on temporary layoff and searching). This equation was estimated with the IWLS technique described in Section I, and the results for the first and fifth (final) iterations are presented in the first two columns of Table 1.

The estimated coefficients on ST_i, the constant term, and CITY, are positive and significant, the latter coefficient being consistent with Seater's theoretical model in which an increased density of potential employers has a favorable effect on contacting firms. Among the interactive terms, only ($UC_i * ST_i$) proves significant, reflecting that, with the receipt of unemployment insurance benefits, an individual also receives assistance from the state agency concerning available jobs. Alternatively, UC_i may segment the sample between those with and without a recent attachment to the labor market, the former presumably having a better knowledge of existing opportunities. In order to determine the marginal returns to search, we examine the estimated coefficient on ST_i^2, which is negative and significantly different from zero. Thus, based on a dataset that reflects direct search activity, we are able to reject the null hypothesis (5) at the 1% level of significance in favor of diminishing returns to search.[4/]

B. Direct + Indirect Search - Sample B

This dataset includes Sample A plus those respondents who contacted firms most frequently using indirect methods. These methods are divided between a self-directed search strategy, where contacts are made as a result of friends or advertisements, and an intermediary strategy, where contacts are made through state or private employment agencies.[5/] In so far as direct and indirect methods may have differing impacts on the relationship between employee contacts and search time, the regression model has been augmented with dummy variables for respondents who most frequently used self-directed (SDS_i) and intermediary (IS_i) search strategies.

The estimates for Sample B, displayed in columns 3 and 4 of Table 1, are broadly similiar to those obtained previously, as all of the coefficients significant in the former sample remain significant. In Sample B, the interactive term ($RACE_i * ST_i$) emerges with a negative and significant coefficient, possibly reflecting the adverse effects of discrimination at employment agencies, or the lack of an effective "network" of friends and relatives for nonwhites.
In either sample, the coefficients on SEX_i and TL_i (expected to be negative due to sexual discrimination and lowered search intensity, respectively), and AGE_i (capturing a positive experience effect) do not prove statistically significant. The negative and significant coefficients on SDS_i and IS_i indicate that indirect methods lead to fewer employee contacts per hour of search time than direct methods. The estimated coefficient on ST_i^2 is smaller is absolute value than before, reflecting that diminishing returns "set-in" slower when using indirect methods, and the null hypothesis is again rejected at the 1% level.

Thus, the empirical evidence presented in this note suggests that the hypothesis of diminishing returns to job search, which underlies a class of theoretical search models and may explain partly the low levels of search by the unemployed, can be supported by the data.

APPENDIX *

In this appendix, the sufficient conditions for which the likelihood function will have a unique maximum are derived. Given the GLS interpretation of the maximum likelihood estimator, critical values for the likelihood function are found by minimizing the following expression.

$$\min_{C} \quad (N - XC)' \, W^{-1} \, (N - XC) \tag{1A}$$

It is more convenient to write (1A) in scalar notation.

$$\min_{\{c_k\}} \sum_{i=1}^{I} \frac{\left(N_i - \sum_{k=1}^{K} c_k x_{i,k}\right)^2}{\sum_{k=1}^{K} c_k x_{i,k}} \tag{2A}$$

where N_i, c_k, and $x_{i,k}$ are elements of N, C, and X, respectively.

A unique maximum to the likelihood function is assured if (2A) is a convex function, which is equivalent to the Hessian of (2A) being positive definite. Differentiating (2A),

$$\frac{\delta}{\delta c_k} = \sum_{i=1}^{I} x_{i,k}\left(1 - \left(N_i \Big/ \sum_{k=1}^{K} c_k x_{i,k}\right)^2\right) \qquad k=1,K \tag{3A}$$

Differentiating (3A),

$$\frac{\delta^2}{\delta c_k \delta c_j} = \sum_{i=1}^{I} (2\, N_i^2\, x_{i,k}\, x_{i,j}) \Big/ \left(\sum_{k=1}^{K} c_k x_{i,k}\right)^3 \qquad k=1,K \quad j=1,K \tag{4A}$$

Letting

$$\gamma_i = 2\, N_i^2 \Big/ \left(\sum_{k=1}^{K} c_k x_{i,k}\right)^3 \qquad i=1,I \tag{5A}$$

(4A) can be written

$$\frac{\delta^2}{\delta c_k \delta c_j} = \sum_{i=1}^{I} \gamma_i\, x_{i,k}\, x_{i,j} \qquad k=1,K \quad j=1,K \tag{6A}$$

---

* The paper to be published in The American Economic Review will not contain this Appendix.

which are the elements of the Hessian

$$H = \begin{bmatrix} \sum_{i=1}^{I} \gamma_i \, x_{i,k} \, x_{i,j} \\ k=1,K \\ j=1,K \end{bmatrix} \qquad (7A)$$

Defining

$$H_i = \begin{bmatrix} \gamma_i \, x_{i,k} \, x_{i,j} \\ k=1,K \\ j=1,K \end{bmatrix} = \gamma_i \begin{bmatrix} x_{i,k} \, x_{i,j} \\ k=1,K \\ j=1,K \end{bmatrix} \quad i=1,I \qquad (8A)$$

H can be rewritten by summing H_i across all i's.

$$H = \sum_{i=1}^{I} H_i \qquad (9A)$$

It is straightforward to show that the summation of positive definite (PD) and positive semi-definite (PSD) matrices is itself PD. In order to establish that H is PD, it will be sufficient to show that, for an arbitrary i, H_i is PSD and, for some i, H_i is PD. (8A) can be written as

$$H_i = \gamma_i \, X^{i\prime} X^i \qquad i=1,I \qquad (10A)$$

where $X^i = [x_{i,1}, x_{i,2}, \ldots, x_{i,k}]$

(10A) will be PSD if $\gamma_i > 0$ and $X^{i\prime}X^i$ is PSD.

From (5A), a sufficient condition for γ_i to be positive is

$$\sum_{k=1}^{K} c_k x_{i,k} = \hat{N}_i > 0 \qquad i=1,I \qquad (11A)$$

$X^{i\prime}X^i$ will be PSD if

$$A' X^{i\prime} X^i A \geq 0 \qquad (12A)$$

where A is a Kx1 vector, not all of whose elements are zero. By matrix manipulation, (12A) becomes

$$(X^i A)' \, (X^i A) \geq 0 \qquad (13A)$$

Since X^iA is a scalar, $(X^iA)' = X^iA$, (13A) can be written

$$(X^i A)^2 \geq 0 \qquad (14A)$$

Thus, $X^{i\prime}X^i$ is PSD for an arbitrary i.

However, under standard conditions of the GLS model, it can not be the case that $X^{i\prime}X^i$ is PSD for all i. Suppose $X^iA = 0$ for i=1,I. Then we can write the following expression.

$$a_1X_1 + a_2X_2 + \ldots + a_KX_K = \theta \tag{15A}$$

where $a_j \varepsilon A$ j=1,K
X_j = Ix1 vector of explanatory variables, j=1,K
$X_j \varepsilon X$
θ = Kx1 vector of zeros.

(15A) implies that $\{X_j\} \varepsilon X$ are linearly dependent and X is of rank less than K. $X'W^{-1}X$ is of rank K by an assumption of the GLS model and the empirical fact that $(X'W^{-1}X)^{-1}$ can be computed. Furthermore, the rank of $X'W^{-1}X$ is equal to the rank of X. (The lemmas cited in the following demonstration are found in Theil (1971), pp. 11-12.)

Let

$$W^* = \begin{bmatrix} \hat{N}_1^{\frac{1}{2}} & & & \\ & \hat{N}_2^{\frac{1}{2}} & & \\ & & \ddots & \\ & & & \hat{N}_I^{\frac{1}{2}} \end{bmatrix} \tag{16A}$$

$$X^* = W^*X \tag{17A}$$

then, by Lemma B.4,

$$\text{rank}(X'W^{-1}X) = \text{rank}(X^{*\prime}X^*) = \text{rank}(X^*) \tag{18A}$$

By Lemma C.3, $\hat{N}_i > 0$, i=1,I, implies $\det(W^*) \neq 0$. Thus, W^* is non-singular and, using Lemma C.6, we can conclude

$$\text{rank}(X^*) = \text{rank}(W^*X) = \text{rank}(X) \tag{19A}$$

Therefore, $X^iA = 0$, for all i=1,I, leads to a contradiction, and we conclude that H_i is PD for at least one i.

Therefore, provided $X'W^{-1}X$ is of full rank and $\hat{N}_i > 0$, i=1,I, (2A) is strictly convex and we are assured that the likelihood function possesses a unique maximum.

REFERENCES

J. M. Barron and O. W. Gilley, "Job Search and Vacancy Contacts: Note," American Economic Review, September 1981, 71, 747-752.

J. M. Barron and W. Mellow, "Search Effort in the Labor Market," The Journal of Human Resources, Summer 1979, 14, 389-404.

R. S. Chirinko, "An Empirical Investigation of the Returns to Job Search," Discussion Paper No. 452, Center for Mathematical Studies in Economics and Management Science, October 1980.

C. L. Gilbert, "Econometric Models for Discrete Economic Processes," Paper read at the Econometric Society European Meeting, Athens, September 3-6, 1979.

M. Nerlove, "Recent Empirical Studies of the CES and Related Production Functions," in The Theory and Empirical Analysis of Production, ed. by Murray Brown, New York,1967.

C. Rosenfeld, "Job Search of the Unemployed, May 1976," Monthly Labor Review, November 1977, 100, 39-43; reprinted as Special Labor Force Report 210, Bureau of Labor Statistics, U.S. Department of Labor.

J. Seater, "Job Search and Vacancy Contacts," American Economic Review, June 1979, 69, 411-419.

R. J. Willis and S. Rosen, "Education and Self-Selection," Journal of Political Economy, October 1979 (part II), 87, 7-36.

TABLE 1

Iterated Weighted Least Squares (IWLS) Estimates
of the Contacts-Search Time Model

Variable	Direct Search		Direct + Indirect Search	
	First Iteration	Fifth Iteration	First Iteration	Fifth Iteration
	(1)	(2)	(3)	(4)
Constant	3.256[a]	3.284[a]	4.642[a]	4.581[a]
	(.6769)	(.5132)	(.5091)	(.4229)
ST_i	.2196[a]	.2623[a]	.1612[a]	.1845[a]
	(.0388)	(.0418)	(.0255)	(.0273)
ST_i^2	-.9998*[a]	-1.192*[a]	-.6539*[a]	-.7714*[a]
	(.2276)*	(.2463)*	(.1494)*	(.1608)*
$AGE_i * ST_i$	.0196*	-.3733*	.4383*	.2343*
	(.6707)*	(.8257)*	(.4541)*	(.5554)*
$RACE_i * ST_i$	-.0197	-.0297	-.0192	-.0282[c]
	(.0189)	(.0222)	(.0129)	(.0147)
$SEX_i * ST_i$	.0109	.0078	-.0128	-.0161
	(.0184)	(.0221)	(.0118)	(.0134)
$UC_i * ST_i$	.0580[a]	.0596[a]	.0481[a]	.0468[a]
	(.0153)	(.0195)	(.0107)	(.0134)
$TL_i * ST_i$	.0023	.0040	.0192	.0178
	(.0256)	(.0323)	(.0184)	(.0235)
$CITY_i$	2.037[a]	1.207[b]	1.338[a]	.6976[c]
	(.6841)	(.5794)	(.4782)	(.3972)
SDS_i	---	---	-2.667[a]	-2.159[a]
	---	---	(.4952)	(.4163)
IS_i	---	---	-2.857[a]	-1.940[a]
	---	---	(.6826)	(.5782)
# of Cases	835	835	1647	1647

NOTE: Estimated standard errors in parentheses.

* Multiplied by 10^{-3}.

a Significant at the 1% level

b Significant at the 5% level

c Significant at the 10% level

FOOTNOTES

1. The likelihood function is constructed under the assumption of independently distributed N_i's. The inclusion of explanatory variables to account for differing abilities in obtaining contacts due to demographic and labor market characteristics will tend to preserve the independence of the N_i's.

2. Sufficient conditions under which the likelihood function will possess a unique maximum, hence a unique $\hat{B}$, are $(X'W^{-1}X)$ being of full rank and $\hat{N}_i > 0$ for all i. The derivation of these conditions, which are satisfied by the estimates presented in this note, can be found in Chirinko (1980).

3. For a further discussion of the survey, see Rosenfeld (1977).

4. Econometric problems remain due to the endogeneity of ST_i and the composition of the sample. The residuals from a cross section regression will tend to represent individual characteristics influencing the amount of time devoted to search, determined by optimizing behavior with regard to a utility function and binding constraints (cf. Nerlove (1967), p. 107). Thus there may exist a correlation between the error term and ST_i resulting in biased coefficients. However, the inclusion in the regression of demographic and labor market variables, which affect the optimal level of search time, will attenuate this problem. Since low efficiency searchers may have a low probability of "escaping" unemployment, the estimates may be tainted by a sample selection bias. Correcting for the non-randomness of the sample may be particularly difficult for computational reasons, owing to the presence of many different categories within the sample (cf. the present situation, where categories are defined by the number of weeks unemployed, to the dichotomous classification of Willis and Rosen (1979)).

5. Differentiating between direct and indirect, and self-directed and intermediary search methods has been suggested by Barron and Gilley (1981).

DISCUSSION PAPER NO. 452

AN EMPIRICAL INVESTIGATION

OF THE RETURNS TO JOB SEARCH

Robert S. Chirinko*

October 1980

* Northwestern University, and Wharton Econometric Forecasting Associates.

I wish to acknowledge the extensive comments and assistance offered by Stephen R. Cosslett, Walter D. Fisher, Shelly J. Lundberg, Wesley Mellow, Dale T. Mortensen, Edward C. Prescott, and participants in the Macro/Labor Workshop at Northwestern University. All errors and omissions remain the sole responsibility of the author.

October 1980

AN EMPIRICAL INVESTIGATION OF THE RETURNS TO JOB SEARCH

ABSTRACT

The assumption of diminishing returns to job search is crucial to the results derived from a number of theoretical search models. This assumption is tested with an econometric model that is general enough to allow for decreasing, constant, or increasing returns to search.

The model is estimated by a recently developed Iterated Weighted Least Squares technique which, under certain conditions, is equivalent to maximum likelihood. Sufficient conditions under which the estimator will be unique are derived. The estimator converges quickly to a unique, maximum likelihood solution, and the results suggest that the hypothesis of diminishing returns to search can be supported by the data.

AN EMPIRICAL INVESTIGATION OF THE RETURNS TO JOB SEARCH

The assumption of diminishing returns to job search plays a crucial role in the results derived from a search model presented recently by John Seater (1979). This assumption is not uncommon in search models but Seater's paper differs in that he motivates this assumption by noting that "spatial aspects of the search process cause diminishing returns." The key idea behind this result is that the travel time required to search firms in a circular area increases faster than the area to be searched.

In this paper, we examine whether the assumption of diminishing returns to job search can be supported by the data. The point of departure is to measure the returns to search by the number of firms that an individual contacts for a given amount of search time. We do not measure the returns to search by the number of job offers obtained because these offers will be influenced strongly by the existing vacancies, a variable that is not under the control of the searcher and will vary widely over time and space. The emphasis on the number of firms contacted is in keeping with the notions developed in the Seater paper, for his model generates diminishing returns as a result of restrictions in the contact "technology," with vacancies entering as a multiplicative constant. Thus, we focus on the ability of individuals to translate search time into job contacts, and evaluate the returns to search by noting whether this "production process" exhibits diminishing returns with respect to the input variable.

Estimating the relationship between firms contacted and search time by Ordinary Least Squares (OLS) is inappropriate since the former variable is discrete, while the latter is continuous. In order to avoid this problem, a two-step procedure is utilized. First, the number of firms contacted by an individual is modeled as a Poisson process. In turn, the Poisson parameter, λ, is related to search time and a set of dummy variables that capture demographic and labor market effects. The model is estimated by Iterated Weighted Least Squares (IWLS) which, in the case of a Poisson process, is equivalent to a maximum likelihood estimator.[1/] Sufficient conditions for the existence of a unique maximum to the likelihood function are derived. The model is estimated using two different samples, and the evidence rejects strongly the hypothesis of constant in favor of diminishing returns to job search.

I. ECONOMETRIC MODEL

In this section, we develop an econometric model in a two-step procedure that allows for an estimation of the technology that transforms search time into employer contacts. First, in order to account for the continuous/discrete nature of this relationship, the number of firms contacted in a fixed period of time is assumed to "arrive" at a rate described by a Poisson distribution.

$$\text{PROB}(N_i = K \mid \lambda_i) = e^{-\lambda_i} \lambda_i^K / K! \qquad (1)$$

N_i = The number of firms contacted in a given period by i.
K = 0,1,2,...
λ_i = Expected number of firms that will be contacted by i (i.e., $E(N_i) = \lambda_i$, $E(\cdot)$ is the expected value operator).
i = 1,I (individuals).

In turn, the Poisson parameter for individual i, λ_i, is related to search time,and a set of demographic and labor market variables that are multiplied by search time. These interactive terms capture those factors that might influence significantly the frequency with which an individual contacts firms, for a given amount of search time. In order to allow for the possibility of non-constant returns to search, a term squared in search time is included as an explanatory variable.

$$\lambda_i = E(N_i) = a + b_1 * ST_i + b_2 * ST_i^2 + \sum_{j=1}^{J} b_{j+2} * ST_i * z_{i,j} \tag{2}$$

ST_i = The total hours spent looking for work in a given period by i.

$z_{i,j}$ = The jth demographic or labor market variable.

$a, b_1, \ldots, b_{J+2}$ = Parameters to be estimated.

The model, equations (1) and (2), can be estimated by the method of maximum likelihood. Assuming that the N_i's are independent,[2/] the log-likelihood function can be written as

$$L(N|\lambda) = \sum_{i=1}^{I} (-\lambda_i + N_i * Ln(\lambda_i) - Ln(N_i!)) \tag{3}$$

Note that λ_i is a function of ST_i. Gilbert (1979) has shown that the maximum likelihood estimator for a Poisson process is equivalent to a Generalized Least Squares estimator (GLS) that can be calculated easily with conventional software packages.

The notation for this model can be generalized in the following manner.

$$\lambda_i = X^i C \qquad i=1,I \tag{4}$$

$C = (a: b_1: b_2: \ldots: b_{J+2})$ - Kx1 vector of parameters.

$X^i = (1: ST_i: ST_i^2: ST_i * z_{i,1}: \ldots: ST_i * z_{i,J})$

- 1xK vector of explanatory variables for the ith unemployed job seeker.

For all individuals, (4) can be written

$$\lambda = X C \tag{5}$$

λ = Ix1 vector of λ_i's.
X = IxK matrix of explanatory variables.

With the model written compactly as (3) and (5), Gilbert's GLS estimator for C is

$$\hat{C} = (X'W^{-1}X)^{-1}X'W^{-1}N \tag{6}$$

N = Ix1 vector of dependent variables.

W = IxI diagonal weighting matrix defined as follows.

$$W = \begin{bmatrix} X^1\hat{C} & & & & & & \\ & X^2\hat{C} & & & & & \\ & & X^3\hat{C} & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & X^I\hat{C} \end{bmatrix}$$

OR (7)

$$W = \begin{bmatrix} \hat{N}_1 & & & & & & \\ & \hat{N}_2 & & & & & \\ & & \hat{N}_3 & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ & & & & & \cdot & \\ & & & & & & \hat{N}_I \end{bmatrix}$$

X^i = The ith row of X. $\qquad$ i=1.I

The estimator given by (6) and (7) is computed iteratively. First, (6) is estimated with W equal to the identity matrix, thus a standard OLS regression is performed. These first-round estimates of $\hat{C}$ are inserted into (7) and a GLS estimation of (6) is implemented. The process of recomputing W and reestimating $\hat{C}$ continues until the change in particular values of $\hat{C}$ from iteration to iteration is less than some predetermined amount.

Given $\hat{C}$, Gilbert has shown that the asymptotic covariance matrix for $\hat{C}$ is given by

$$V(\hat{C}) = (X'W^{-1}X)^{-1} \tag{8}$$

Since the model is being estimated by a maximum likelihood technique, we need some conditions to ensure that $\hat{C}$ is the unique solution to the likelihood function (3). In the appendix, we show that $X'W^{-1}X$ being of full rank and the estimated number of contacts, $\hat{N}_i$, being positive are sufficient conditions for the uniqueness of $\hat{C}$.

The returns to search hypothesis is stated in terms of the coefficient on ST_i^2.

$$\begin{aligned} H_o &: b_2 = 0.0 \\ H_a &: b_2 \neq 0.0 \end{aligned} \tag{9}$$

Given a large sample, $\hat{C}$, and $V(\hat{C})$, (9) can be evaluated, using a normal distribution, to determine if the null hypothesis can be rejected in favor of the alternative.

II. EMPIRICAL RESULTS

The model is estimated using two samples drawn from a special survey of unemployed respondents in the Current Population Survey (CPS), May 1976.[3/] For the previous four week period the respondents were asked, among other questions, the number of hours spent looking for work, the method used most frequently, and the number of firms contacted. Those respondents who classified themselves as being on layoff and not looking for another job, or unemployed and not spending any time searching are excluded from the samples used in this paper. Since participants in the special survey were also included in the CPS, we are able to assess the influence on the unemployed of a broad range of demographic and labor market characteristics.

Before examining the empirical results, we discuss two biases that may affect the estimated coefficients. First, a frequent problem in estimating conventional, value-added production functions, with output specified as some function of the inputs, is the possibility of a significant correlation between the omitted variables, as represented by the regression residuals, and the choice variables of the decision making unit.[4/] In the cross section analysis of this paper, the residuals will tend to represent differences among individuals that will influence the amount of time devoted to search, as they optimize with regard to their utility functions and associated constraints. Thus there may exist a correlation between the error term and explanatory variables in the regression model, and the estimated coefficients may be biased. Using the search model of Seater as a point of reference, we are unable to determine analytically the direction of the bias on the estimated coefficients since "the response of search [to an increase in the returns to search] is ambiguous because the income and substitution effects on search oppose each other (Seater (1977), p. 361)." However, the inclusion in the regression equation of variables representing demographic and labor market characteristics, which influence the optimal level of search time, will tend to minimize this bias.

Second, the sample may be contaminated by a self-selection bias. Search models imply that individuals will choose all work or all search if they face constant returns to search (Seater (1979), p. 415). If this model is "correct," then since the dataset is comprised entirely of unemployed job searchers, the sample will not be representative of the population, and the estimated coefficient on the square of search time will be biased towards zero. This bias will not affect the overall conclusion of this paper, however, if the null hypothesis is rejected despite the biased coefficient.

A. Direct Search - Sample A

This sample is restricted to those respondents whose most frequent method of search was to "apply directly to employers without suggestions or referrals by anyone," and it is this type of search activity that is closest to Seater's notion of spatial search. The regression equation is specified with the number of firms contacted by an individual dependent on search time (ST_i), the square of search time, a constant, a dummy variable, CITY, which takes a value of 1 if the respondent lives in a SMSA,[5/] and interactive terms, defined as search time multiplied separately by a series of variables: AGE_i (equal to chronological age), $RACE_i$ (dummy variable, 1 if nonwhite), SEX_i (dummy variable, 1 if female), UC_i (dummy variable, 1 if receiving unemployment insurance benefits), and TL_i (dummy variable, 1 if on temporary layoff and searching). This equation was estimated with the IWLS technique described in Section I, and the results for the first and fifth (final) iterations are presented in the first two columns of Table 1.

The estimated coefficients on ST_i and the constant term are both positive and significant at the 1% level, the significant constant indicating that some relevant explanatory variables may have been omitted from the regression equation. The coefficient on CITY is also positive, and significant at the 5% level, reflecting that an increased density of potential employers has a favorable effect on contacting firms. Among the interactive terms, only $(UC_i * ST_i)$ proves statistically significant. This positive coefficient reflects that, with the receipt of unemployment insurance benefits, an individual also receives information from the state agency concerning available jobs. Furthermore, this variable also serves to segment the sample between those with and without a recent attachment to the labor market, the former presumably having a better knowledge of existing opportunities.

In order to determine the marginal returns to job search, we use the estimated coefficient on ST_i^2, which is negative and significantly different than zero. Thus, based on a dataset that reflects direct search activity, we are able to reject the null hypothesis (9) at the 1% level of significance in favor of the alternative of diminishing returns to search.[6/]

B. Direct + Indirect Search - Sample B

In this sample, the dataset is expanded to include Sample A plus those respondents who contacted firms most frequently using indirect methods. These methods are divided between a self-directed search strategy, where contacts are made as a result of friends or advertisements, and an intermediary strategy, where contacts are made through state or private employment agencies.[7/] In so far as direct and indirect methods may have differing impacts on the contacts-search time relationship, the regression model used above has been augmented with dummy variables for respondents who most frequently used self-directed (SDS_i) or intermediary (IS_i) search strategies. The regression results for Sample B are displayed in columns 3 and 4 of Table 1.

The estimates for Sample B are broadly similiar to those obtained previously. All of the coefficients that were statistically significant in the former sample remain significant and at the same level, with the exception of the coefficient on CITY, which was significant at the 5% level and remains significant, but at the 10% level. Unlike in Sample A, the interactive term ($RACE_i * ST_i$) emerges as significant at the 10% level with a negative coefficient, possibly reflecting the effects of discrimination. The estimated coefficient and standard error on ST_i^2 are both smaller in absolute value than before, and the null hypothesis is again rejected at the 1% level in favor of the alternative of diminishing returns to search.

The negative coefficients on SDS_i and IS_i indicate, somewhat paradoxically, that despite the assistance offered with these search methods, respondents tend to contact fewer firms per hour of search time than those who use direct methods. However, it must be kept in mind that the dependent variable in the regression is the number of firms contacted, but the desideratum of searchers is employment. The probability of employment can be viewed as the product of the probabilities of contacting a firm and, given a contact, of obtaining an acceptable offer. It would then appear quite reasonable for an individual to choose an indirect strategy which, as indicated by the regression results, leads to a lower probability of employer contacts if there is a correspondingly higher probability of obtaining an acceptable job offer.[8/]

It should be noted that the model estimated with either sample converged quickly, and the estimated coefficients and standard errors on ST_i and ST_i^2 all changed by less than .35% between the fourth and fifth iterations. Furthermore, for all iterations and both samples, $(X'W^{-1}X)$ was of full rank, and $\hat{N}_i > 0$, for all i. Thus, using the results developed in the Appendix, we are assured of the uniqueness of the estimates used in rejecting the null hypothesis.

III. SUMMARY

In this paper, the assumption of diminishing returns to job search has been tested with the data by estimating a cross section relationship between the number of firms contacted and the amount of time spent searching by unemployed individuals. In order to account explicitly for the discrete-continuous nature of this relationship, we have employed a two-step procedure that relies on a Poisson process. The equivalence between the maximum likelihood and Iterated Weighted Least Squares techniques facilitated the computations, and the estimator converged quickly to a unique solution. The model has been estimated using two different samples, and the results suggest that the hypothesis of diminishing returns to job search can be supported by the data.

APPENDIX

In this appendix, the sufficient conditions for which the likelihood function will have a unique maximum are derived. Given the GLS interpretation of the maximum likelihood estimator, critical values for the likelihood function are found by minimizing the following expression.

$$\underset{C}{\mathrm{MIN}} \quad (N - XC)' \, W^{-1} \, (N - XC) \tag{1A}$$

It is more convenient to write (1A) in scalar notation.

$$\underset{\{c_k\}}{\mathrm{MIN}} \sum_{i=1}^{I} \frac{\left(N_i - \sum\limits_{k=1}^{K} c_k x_{i,k}\right)^2}{\sum\limits_{k=1}^{K} c_k x_{i,k}} \tag{2A}$$

where N_i, c_k, and $x_{i,k}$ are elements of N, C, and X, respectively.

A unique maximum to the likelihood function is assured if (2A) is a convex function, which is equivalent to the Hessian of (2A) being positive definite. Differentiating (2A),

$$\frac{\delta}{\delta c_k} = \sum_{i=1}^{I} x_{i,k}\left(1 - \left(N_i \Big/ \sum_{k=1}^{K} c_k x_{i,k}\right)^2\right) \qquad k=1,K \tag{3A}$$

Differentiating (3A),

$$\frac{\delta^2}{\delta c_k \delta c_j} = \sum_{i=1}^{I} \left(2\, N_i^2 \, x_{i,k} \, x_{i,j}\right) \Big/ \left(\sum_{k=1}^{K} c_k x_{i,k}\right)^3 \qquad \begin{array}{l} k=1,K \\ j=1,K \end{array} \tag{4A}$$

Letting

$$\gamma_i = 2\, N_i^2 \Big/ \left(\sum_{k=1}^{K} c_k x_{i,k}\right)^3 \qquad i=1,I \tag{5A}$$

(4A) can be written

$$\frac{\delta^2}{\delta c_k \delta c_j} = \sum_{i=1}^{I} \gamma_i \, x_{i,k} \, x_{i,j} \qquad \begin{array}{l} k=1,K \\ j=1,K \end{array} \tag{6A}$$

which are the elements of the Hessian

$$H = \left[\sum_{i=1}^{I} \gamma_i \, x_{i,k} \, x_{i,j} \right]_{\substack{k=1,K \\ j=1,K}} \tag{7A}$$

Defining

$$H_i = \left[\gamma_i \, x_{i,k} \, x_{i,j} \right]_{\substack{k=1,K \\ j=1,K}} = \gamma_i \left[x_{i,k} \, x_{i,j} \right]_{\substack{k=1,K \\ j=1,K}} \qquad i=1,I \tag{8A}$$

H can be rewritten by summing H_i across all i's.

$$H = \sum_{i=1}^{I} H_i \tag{9A}$$

It is straightforward to show that the summation of positive definite (PD) and positive semi-definite (PSD) matrices is itself PD. In order to establish that H is PD, it will be sufficient to show that, for an arbitrary i, H_i is PSD and, for some i, H_i is PD. (8A) can be written as

$$H_i = \gamma_i \, X^{i\prime} X^i \qquad i=1,I \tag{10A}$$

$$\text{where } X^i = [x_{i,1}, x_{i,2}, \ldots, x_{i,k}]$$

(10A) will be PSD if $\gamma_i > 0$ and $X^{i\prime}X^i$ is PSD.

From (5A), a sufficient condition for γ_i to be positive is

$$\sum_{k=1}^{K} c_k x_{i,k} = \hat{N}_i > 0 \qquad i=1,I \tag{11A}$$

$X^{i\prime}X^i$ will be PSD if

$$A' X^{i\prime} X^i A \geq 0 \tag{12A}$$

where A is a Kx1 vector, not all of whose elements are zero. By matrix manipulation, (12A) becomes

$$(X^i A)' \, (X^i A) \geq 0 \tag{13A}$$

Since X^iA is a scalar, $(X^iA)' = X^iA$, (13A) can be written

$$(X^i A)^2 \geq 0 \tag{14A}$$

Thus, $X^{i}{}'X^{i}$ is PSD for an arbitrary i.

However, under standard conditions of the GLS model, it can not be the case that $X^{i}{}'X^{i}$ is PSD for all i. Suppose $X^{i}A = 0$ for i=1,I. Then we can write the following expression.

$$a_1X_1 + a_2X_2 + \ldots + a_KX_K = \theta \tag{15A}$$

where $a_j \in A$ j=i,K

X_j = Ix1 vector of explanatory variables, j=1,K

$X_j \in X$

θ = Kx1 vector of zeros

(15A) implies that $\{X_j\} \in X$ are linearly dependent and X is of rank less than K. $X'W^{-1}X$ is of rank K by an assumption of the GLS model and the empirical fact that $(X'W^{-1}X)^{-1}$ can be computed. Furthermore, the rank of $X'W^{-1}X$ is equal to the rank of X. (The lemmas cited in the following demonstration are found in Theil (1971), pp. 11-12.)

Let

$$W^* = \begin{bmatrix} \hat{N}_1^{\frac{1}{2}} & & & \\ & \hat{N}_2^{\frac{1}{2}} & & \\ & & \ddots & \\ & & & \hat{N}_I^{\frac{1}{2}} \end{bmatrix} \tag{16A}$$

$$X^* = W^*X \tag{17A}$$

then, by Lemma B.4,

$$\text{rank}(X'W^{-1}X) = \text{rank}(X^*{}'X^*) = \text{rank}(X^*) \tag{18A}$$

By Lemma C.3, $\hat{N}_i > 0$, i=1,I, implies $\det(W^*) \neq 0$. Thus, W^* is non-singular and, using Lemma C.6, we can conclude

$$\text{rank}(X^*) = \text{rank}(W^*X) = \text{rank}(X) \tag{19A}$$

Therefore, $X^{i}A = 0$, for all i=1,I, leads to a contradiction, and we conclude that H_i is PD for at least one i.

Therefore, provided $X'W^{-1}X$ is of full rank and $\hat{N}_i > 0$, i=1,I, (2A) is strictly convex and we are assured that the likelihood function possesses a unique maximum.

REFERENCES

Barron, John M., and Gilley, Otis W. "Job Search and Vacancy Contacts: A Comment." Unpublished Manuscript, December 1979.

Gilbert, C. L. "Econometric Models for Discrete Economic Processes." Paper read at the Econometric Society European Meeting, Athens, September 3-6, 1979.

Nerlove, Marc. "Recent Empirical Studies of the CES and Related Production Functions." In _The Theory and Empirical Analysis of Production_, edited by Murray Brown. New York: Columbia University Press for the National Bureau of Economic Research, 1967.

Rosenfeld, Carl. "Job Search of the Unemployed, May 1976." _Monthly Labor Review_ 100 (November 1977): 39-43; reprinted as Special Labor Force Report 210, Bureau of Labor Statistics, U.S. Department of Labor.

Seater, John J. "A Unified Model of Consumption, Labor Supply, and Job Search." _Journal of Economic Theory_ 14 (April 1977): 349-372.

_______. "Job Search and Vacancy Contacts." _American Economic Review_ 69 (June 1979): 411-419.

Theil, Henri. _Principles of Econometrics_ New York: John Wiley and Sons, 1971.

FOOTNOTES

1. This equivalence has been demonstrated by Gilbert (1979), pp. 4-5.

2. The inclusion of variables to account for differing abilities in obtaining contacts due to demographic or labor market characteristics will tend to preserve the independence of the N_i's.

3. For a further discussion of the survey, see Rosenfeld (1977).

4. For a discussion of this problem in regards to conventional, value-added production functions, see Nerlove (1967), especially p. 107.

5. All dummy variables used in this study assume integer values of either 0 or 1.

6. Utilizing a similiar dataset and a linear-in-logs specification, Barron and Gilley (1979) also found evidence supporting diminishing returns to search. A linear-in-levels version of the Barron-Gilley model was reestimated using the IWLS technique, and we were able to reject the null hypothesis in favor of the alternative of diminishing returns. However, when compared to the converged IWLS solution, the estimated standard errors on the coefficient relevant to the hypothesis test was higher by 57% in the linear-in-logs (original Barron-Gilley) and by 25% in the linear-in-levels (OLS) versions of the model.

7. Differentiating between direct and indirect, and self-directed and intermediary search methods has been suggested by Barron and Gilley (1979).

8. This argument has been put forth by Barron and Gilley (1979), and they discuss empirical evidence supporting this view.

TABLE 1

Iterated Weighted Least Squares (IWLS) Estimates of the Contacts-Search Time Model

Variable	Direct Search: First Iteration	Direct Search: Fifth Iteration	Direct + Indirect Search: First Iteration	Direct + Indirect Search: Fifth Iteration
Constant	3.256 (.6769)	3.284 (.5132)	4.642 (.5091)	4.581 (.4229)
ST_i	.2196 (.0388)	.2623 (.0418)	.1612 (.0255)	.1845 (.0273)
ST_i^2	-.9998* (.2276)*	-1.192* (.2463)*	-.6539* (.1494)*	-.7714* (.1608)*
$AGE_i * ST_i$	.0196* (.6707)*	-.3733* (.8257)*	.4383* (.4541)*	.2343* (.5554)*
$RACE_i * ST_i$	-.0197 (.0189)	-.0297 (.0222)	-.0192 (.0129)	-.0282 (.0147)
$SEX_i * ST_i$	.0109 (.0184)	.0078 (.0221)	-.0128 (.0118)	-.0161 (.0134)
$UC_i * ST_i$	.0580 (.0153)	.0596 (.0195)	.0481 (.0107)	.0468 (.0134)
$TL_i * ST_i$	2.285* (25.64)*	3.972* (32.31)*	.0192 (.0184)	.0178 (.0235)
$CITY_i$	2.037 (.6841)	1.207 (.5794)	1.338 (.4782)	.6976 (.3972)
SDS_i	--- ---	--- ---	-2.667 (.4952)	-2.159 (.4163)
IS_i	--- ---	--- ---	-2.857 (.6826)	-1.940 (.5782)
# of Cases	835	835	1647	1647

NOTE: Estimated standard errors in parentheses.

* Multiplied by 10^{-3}.