

- and *Labor Relations Review* 31 (Oct. 1977), 45-60.
- \_\_\_\_\_, "Concentration and Professional Earnings in Manufacturing," *Industrial and Labor Relations Review* 31 (Apr. 1978), 379-384.
- Freeman, Richard B., and James L. Medoff, "The Impact of the Percentage Organized on Union and Nonunion Wages," this REVIEW 63 (Nov. 1981), 561-572.
- Haworth, Charles T., and David W. Rasmussen, "Human Capital and Inter-Industry Wages in Manufacturing," this REVIEW 53 (Nov. 1971), 376-380.
- Haworth, Charles T., and Carol J. Reuther, "Industrial Concentration and Inter-Industry Wage Determination," this REVIEW 60 (Feb. 1978), 85-95.
- Hendricks, Wallace, "Unionism, Oligopoly and Rigid Wages," this REVIEW 63 (May 1981), 198-205.
- Levinson, Harold M., "Unionism, Concentration and Wage Changes: Toward a Unified Theory," *Industrial and Labor Relations Review* 20 (Jan. 1967), 198-205.
- Masters, S. H., "Wages and Plant Size: An Interindustry Analysis," this REVIEW 51 (Aug. 1969), 341-345.
- Mellow, Wesley, "Employer Size and Wages," this REVIEW 64 (Aug. 1982), 495-501.
- Miller, Edward M., "Large Firms Are Good For Their Workers: Manufacturing Wages as a Function of Firm Size and Concentration," *The Antitrust Bulletin* (Spring 1981), 145-154.
- Pugel, Thomas A., "Profitability, Concentration and the Interindustry Variation in Wages," this REVIEW 62 (May 1980), 248-253.
- Rapping, Leonard A., "Monopoly Rents, Wage Rates, and Union Wage Effectiveness," *Quarterly Review of Economics and Business* 7 (Spring 1967), 31-47.
- Ross, Stephen A., and Michael L. Wachter, "Wage Determination, Inflation, and the Industrial Structure," *The American Economic Review* 63 (Sept. 1973), 675-692.
- Segal, Martin, "The Relation between Union Wage Impact and Market Structure," *Quarterly Journal of Economics* 78 (Feb. 1964), 96-114.
- Sparrow, Dorothy, "Changing Market Structure: Impact on Union Wage Gains," *Industrial Relations* 10 (May 1971), 160-167.
- Stafford, Frank P., "Concentration and Labor Earnings: Comment," *The American Economic Review* 58 (Mar. 1968), 174-181.
- Weiss, Leonard W., "Concentration and Labor Earnings," *American Economic Review* 56 (Mar. 1966), 96-117.

## A LACK-OF-FIT TEST FOR ECONOMETRIC APPLICATIONS TO CROSS-SECTION DATA

Howard E. Doran and Jan Kmenta\*

*Abstract*—A lack-of-fit test of model specification used by experimental statisticians but mostly unknown to econometricians is presented. The test is applicable in situations in which there are replicated observations on the dependent variable. In this paper the test is modified to allow for heteroskedasticity usually encountered when dealing with cross-sectional observations, and illustrated by an application to an earnings function estimated from a sample survey of Norwegian women.

### I. Introduction

The lack-of-fit (LOF) test is a model specification test that can be applied when there are replicated observations on the dependent variable corresponding to observations on the explanatory variables. Basically, the test utilizes the additional information that comes from

within-group variation. As the situation of replicated observations is normal in experimental work, the test is well known to experimental statisticians but appears to be almost unknown to econometricians. (A survey of econometric text books has revealed no mention of the test.) This is probably due to the strong emphasis in econometrics on the methodology applicable to time-series data involving a *single* observation on the dependent variable for each set of observed values of the explanatory variables. By contrast, when cross-sectional data are used, there can be many units (individuals, firms, families) that are characterized by the same values of explanatory variables (e.g., incomes, prices, educational levels). In this situation a lack-of-fit test could often be profitably applied as an aid to appropriate model specification.

The main purpose of the paper is to draw the attention of econometricians to the possibilities offered by the lack-of-fit test (see also Battese (1977)). The paper is organized as follows. Section II contains a development and an explanation of the test. In section III the test is generalized to allow for heteroskedasticity which frequently characterizes relations pertaining to cross-sectional observations. Finally, in section IV the test is applied to the standard semilog earnings function due to Mincer (1974), utilizing data from the Norwegian Fertility Survey of 1977.

Received for publication September 24, 1984. Revision accepted for publication June 10, 1985.

\* University of New England, Australia, and University of Michigan, respectively.

A major part of this work was done while the authors were Fellows of the Alexander von Humboldt Foundation at the University of Bonn, Germany. Comments by Michael McAleer, Saul Hymans, Karl Lin, Jerry Thursby, G. S. Maddala and by the anonymous referees are gratefully acknowledged. Special thanks are due to Helge Brunborg and the Norwegian Central Bureau of Statistics for providing the survey data used in this paper, and to Jeffrey Pliskin who was very helpful with suggestions and who carried out all of the computer work involved.

**II. The Lack-of-Fit Test**

Suppose that an  $n \times 1$  random vector  $Y$  is normally distributed with mean  $\mu$  and covariance matrix  $\sigma^2 I_n$ . The null hypothesis specifies a linear model of the form

$$\mu = X\beta, \tag{1}$$

where  $X$  and  $\beta$  are  $n \times K$  and  $K \times 1$ , respectively. Let us suppose that there are  $m$  ( $m > K$ ) distinct observations on the explanatory variables  $X$ , and that corresponding to the  $i^{\text{th}}$  such observation there are  $n_i$  observations on the dependent variable  $Y$ , where  $n_i \geq 1$  and  $\sum_{i=1}^m n_i = n$ . We will refer to these  $n_i$  observations as the " $i^{\text{th}}$  group." The alternative hypothesis is

$$\mu = X\beta + Z\gamma, \tag{2}$$

where  $Z$  is a matrix of values of *unspecified* omitted explanatory variables (including possibly higher powers of  $X$ ) of dimension  $n \times L$  ( $L < m - K$ ) such that  $Z'X \neq 0$ , and  $\gamma \neq 0$ . Note that we assume that if the model is misspecified, there are  $m$  distinct observations on  $X$  and  $Z$  (although those on  $Z$  are not measured because they are not known), and that for each such observation there are  $n_i$  observations on the dependent variable  $Y$ . That is, under  $H_0$  we have

$$Y_{ij} = x_i\beta + u_{ij} \quad (i = 1, 2, \dots, m; j = 1, 2, \dots, n_i)$$

whereas under  $H_A$  we have

$$Y_{ij} = x_i\beta + z_i\gamma + u_{ij}$$

where  $x_i$  and  $z_i$  represent the  $i^{\text{th}}$  row of  $X$  and  $Z$ , respectively.<sup>1</sup> Both variables  $X$  and  $Z$  are considered to be fixed in repeated samples.

The error sum of squares ( $SSE$ ) from the regression of  $Y$  on  $X$  is given by

$$SSE = \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_{ij})^2, \tag{3}$$

where  $Y_{ij}$  denotes the  $j^{\text{th}}$  observation on the dependent variable in the  $i^{\text{th}}$  group. As the  $i^{\text{th}}$  group observations are all characterized by the same observation on the explanatory variables, the fitted values  $\hat{Y}_{ij}$  ( $j = 1, 2, \dots, n_i$ ) must all be equal. Therefore we may write

$$\hat{Y}_{ij} = \hat{Y}_i,$$

and it follows that

$$\begin{aligned} SSE &= \sum_{i=1}^m \sum_{j=1}^{n_i} [(Y_{ij} - \bar{Y}_i) + (\bar{Y}_i - \hat{Y}_i)]^2 \\ &= \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 + \sum_{i=1}^m n_i (\bar{Y}_i - \hat{Y}_i)^2 \end{aligned}$$

<sup>1</sup> If instead of  $z_i$  the observations on  $Z$  are given by  $z_{ij}$ , the distribution of the test statistic in (11) is unchanged under  $H_0$  and the test is still valid, but the power of the test is adversely affected. In this situation the test is in the class of the Goldfeld-Quandt test when used as a specification error test. We are indebted to Jerry Thursby for a lengthy comment on this point.

where  $\bar{Y}_i$  is a sample mean of the  $i^{\text{th}}$  group. Thus the error sum of squares can be partitioned into two components

$$SSP = \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2 \tag{4}$$

and

$$SSL = \sum_{i=1}^m n_i (\bar{Y}_i - \hat{Y}_i)^2. \tag{5}$$

The first of these components represents the "within-group" variation which, following the experimental literature, we call the "pure error sum of squares" ( $SSP$ ). The second component is termed the "lack-of-fit sum of squares" ( $SSL$ ). It is the error sum of squares that would be obtained if each group were replaced by its sample mean and these sample means regressed on the same regressor variables with each observation weighted by  $\sqrt{n_i}$ . Defining

$$s_i^2 = \frac{1}{(n_i - 1)} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2, \tag{6}$$

we may write

$$SSP = \sum_{i=1}^m (n_i - 1) s_i^2 \tag{7}$$

and  $SSP/(n - m)$  is seen to be a weighted average of  $m$  different estimates  $s_i^2$  of  $\sigma^2$ . This means that we have two different estimators of  $\sigma^2$ ,  $SSE/(n - K)$  and  $SSP/(n - m)$ . Under  $H_0$

$$E[SSE/(n - K)] = E[SSP/(n - m)] = \sigma^2$$

whereas under  $H_A$

$$E[SSE/(n - K)] > E[SSP/(n - m)] = \sigma^2,$$

since variation around any point other than the mean always exceeds the variation around the mean. This provides the basis for the lack-of-fit test.<sup>2</sup>

Now, under  $H_0$

$$SSE/\sigma^2 \sim \chi_{n-K}^2 \tag{8}$$

and

$$SSP/\sigma^2 \sim \chi_{n-m}^2. \tag{9}$$

Further, it can easily be shown that  $SSL (= SSE - SSP)$  and  $SSP$  are independent so that

$$SSL/\sigma^2 = \chi_{m-K}^2, \tag{10}$$

and, therefore, if the model is correctly specified,

$$F = \frac{SSL/(m - K)}{SSP/(n - m)} \sim F_{m-K, n-m}. \tag{11}$$

<sup>2</sup> The paper is directed to situations in which all *primary* replicated data are available. However, if the data are in the form of aggregates such that for each group we are given (i) the group size ( $n_i$ ), (ii) the group mean ( $\bar{Y}_i$ ), and (iii) the group variance ( $s_i^2$ ), then the LOF test can also be applied.

A simple way of calculating *SSP* follows from the fact that *SSP* is simply the error sum of squares obtained by the application of least squares to the unconstrained regression equation

$$Y_{i,j} = \mu_1 D_{1j} + \mu_2 D_{2j} + \dots + \mu_m D_{mj} + u_{i,j} \quad (12)$$

where

$$D_{i,j} = 1 \text{ for all observations in the } i^{\text{th}} \text{ group,} \\ = 0 \text{ otherwise.}$$

Thus *SSE* defined in (3) represents a constrained (by the null hypothesis) error sum of squares whereas *SSP* represents an unconstrained error sum of squares. The *F*-statistic in (11) can equivalently be written as

$$F = \frac{(SSE - SSP)/(m - K)}{SSP/(n - m)} \quad (13)$$

Finally, using the well-known result that  $F_{\nu(1), \nu(2)}$  approaches  $\chi^2_{\nu(1)}/\nu(1)$  as  $\nu(2) \rightarrow \infty$ , we see that as  $n \rightarrow \infty$ ,  $SSL/\sigma^2$  is asymptotically distributed as  $\chi^2_{m-k}$ . In general,  $\sigma^2$  is, of course, unknown. However, if  $\sigma^2$  were known, a more powerful test would be obtained by using the chi-square result (10) rather than (11) or (13). This remark has relevance to the following section.

### III. The Lack-of-Fit Test Under Heteroskedasticity

As mentioned in the introduction, the main area of usefulness in econometrics of the lack-of-fit test is in the analysis of cross-sectional data. In this context the underlying assumption that  $E[(Y - \mu)(Y - \mu)'] = \sigma^2 I_n$  is unlikely to be realistic. We will generalize it by assuming that

$$E[(Y - \mu)(Y - \mu)'] = \Sigma, \quad (14)$$

where

$$\Sigma = \text{diag}(\sigma_1^2 I_{n(1)}, \sigma_2^2 I_{n(2)}, \dots, \sigma_m^2 I_{n(m)})$$

and the  $\sigma_i^2$  are known. Then, defining

$$D = \text{diag}(\sigma_1 I_{n(1)}, \sigma_2 I_{n(2)}, \dots, \sigma_m I_{n(m)}),$$

it follows that

$$D^{-1}y \sim N(D^{-1}\mu, I_n)$$

and the linear specification on the mean under the null hypothesis takes the form

$$D^{-1}\mu = (D^{-1}X)\beta. \quad (15)$$

Thus, provided the analysis is carried out in terms of the weighted observations  $Y_{i,j}/\sigma_i$  and  $X_i/\sigma_i$  ( $i = 1, 2, \dots, m$ ), the test described in the previous section follows through with one difference, namely, that  $\sigma^2$  is known to be equal to unity. Thus by (10), if the model is correctly specified,

$$SSL^* \sim \chi^2_{m-k}, \quad (16)$$

where we have used *SSL\** to emphasize that weighted observations are used.

In practice the  $\sigma_i^2$  are rarely known and have to be estimated. Provided the values of the omitted variables remain unchanged whenever the values of the included variables are unchanged (i.e., provided the model under  $H_A$  is  $Y_{i,j} = x_i\beta + z_i\gamma + u_{i,j}$ ), the estimators  $s_i^2$  defined in (6) are consistent estimators of the  $\sigma_i^2$ . These estimators are completely independent of the specification  $\mu = X\beta$ . Thus the result in (16) is to be regarded as a *large sample* result; if the observations are weighted inversely by  $s_i$ , then as  $n_i \rightarrow \infty$ , the distribution of *SSL\** is given by (16).

It is commonly assumed in econometrics that the variances  $\sigma_i^2$  can be related to a single variable  $w_i$  through a relationship of the form

$$\sigma_i^2 = \alpha w_i^\delta, \quad (w_i > 0), \quad (17)$$

where  $\alpha$  and  $\delta$  are unknown parameters. Typically,  $w_i$  would be a member of the regressor set, but this need not be the case. If such a model is appropriate and  $m > 2$ , then there will be gains in asymptotic efficiency if we make use of the information given in (17). A simple method of using (17) proceeds as follows. Defining  $s_i^2$  as in (6), we have

$$v_i = \frac{(n_i - 1)s_i^2}{\sigma_i^2} \sim \chi^2_{n(i)-1}.$$

Therefore, given that  $\sigma_i^2 = \alpha w_i^\delta$ ,

$$\ln s_i^2 = \ln \alpha + \delta \ln w_i + u_i, \quad (18)$$

where

$$u_i = \ln v_i - \ln(n_i - 1).$$

Bartlett and Kendall (1946) have shown that for large  $n_i$ ,  $\sqrt{(n_i - 2)} u_i$  is approximately normally distributed with mean zero and variance 2. Furthermore, they showed that this approximation is likely to be good for  $n_i$  as small as 10.

Thus least squares regression of  $\sqrt{(n_i - 2)} \ln s_i^2$  on  $\sqrt{(n_i - 2)}$  and  $\sqrt{(n_i - 2)} \ln w_i$  will produce consistent and asymptotically efficient estimates of  $\hat{\alpha}$  and  $\hat{\delta}$ .

### IV. An Illustrative Application of the LOF Test to an Earnings Function for Women

To illustrate the use of the LOF test, we consider the effect of schooling on earnings, and in particular how this effect applies to women whose work history typically differs from that of men. The standard model of earnings proposed by Mincer (1974) and frequently used in applied labor economics—most recently by Behrman and Birdsall (1983) and by Chiswick (1983)—is basically of the form:

$$\ln W_{i,j} = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 S_i + u_{i,j} \quad (19)$$

where

$W_{ij}$  is the wage-rate of the  $j^{\text{th}}$  individual in the  $i^{\text{th}}$  group;

$X_i$  is the “on-job” work experience (common to all individuals in the  $i^{\text{th}}$  group);

$S_i$  is the number of years of schooling (common to all individuals in the  $i^{\text{th}}$  group);

and  $u_{ij}$  is an independently distributed stochastic disturbance. (The subscripts have been adapted to the replicated data case considered by us.) Other variables, such as presence of children or headship of family, that are frequently controlled for in applied work can be subsumed under the omitted  $Z$  variables.

At least two types of possible misspecifications of equation (19) have been discussed in the literature.<sup>3</sup> First, a study by Mincer and Polachek (1974) provides some evidence that the marginal effect of schooling on the wage-rate of people of equal work experience is not constant for women with families. A similar conclusion with respect to heads of households (of any sex) has been reached by Ryder, Stafford, and Stephan (1976) on the basis of a model of life-cycle decision making with leisure as a choice variable. This suggests the possibility of an incorrect functional form of the equation. Second, it is by no means certain that the model in (19) contains all relevant explanatory variables. In particular, Griliches (1977) contends that models such as that in (19) suffer from the fact that individual ability—likely to be correlated with schooling—has been left out. The LOF test can certainly be effective in detecting incorrect functional form since the omitted variables are now represented (at least approximately) by higher powers of  $X$ . As for the omitted “ability” variable, for the test to work well it would have to be true that all individuals of given experience and schooling are characterized by nearly the same ability, otherwise the test will lose some power.<sup>4</sup>

The preceding discussion leads to the conclusion that it may be worthwhile to apply the LOF test not only to equation (19)—to be labeled Model 1—but also to an equation that allows for a nonlinear effect of schooling as in

$$\ln W_{ij} = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \beta_3 S_i + \beta_4 S_i^2 + u_{ij} \quad (20)$$

which we label Model 2. To carry out the test we used the data on about 2,000 married women taking part in

<sup>3</sup> Along with other applied research workers, we do not address the problem of a simultaneous equation bias that may arise from the endogeneity of schooling—a point discussed at length by Griliches (1977).

<sup>4</sup> See footnote 1 above. Griliches (1977) refers to “ability” as “an unobserved latent variable that both drives people to get relatively more schooling and earn more income, given schooling . . .” (p. 7).

the Fertility Survey 1977 carried out by the Norwegian Central Bureau of Statistics. The survey involved a random sample of about 4,000 women but this number was reduced to about one-half after eliminating all observations that were incomplete or that corresponded to a small number of replications. Actually, all that is (theoretically) required is that  $n_i > 1$  for at least one group. The only problem with this arises under heteroskedasticity when  $\sigma_i$  is not known and has to be estimated since this is impossible when  $n_i = 1$ . However, if we make the assumption spelled out in equation (17) above, namely that

$$\sigma_i^2 = \alpha w_i^\delta,$$

then we can use those groups for which  $n_i > K$  to estimate  $\alpha$  and  $\delta$ . Having found  $\hat{\alpha}$  and  $\hat{\delta}$ , we obtain

$$\hat{\sigma}_i^2 = \hat{\alpha} w_i^{\hat{\delta}}$$

which can be applied to *all* groups. From this stage on, all groups are used and no information is thrown away. We did not pursue this refinement because our main purpose was just to illustrate the use of the LOF test. The elimination of the groups with a small number of replications may affect the substantive interpretation—if not the validity—of our results.

With respect to measurement of the variables involved, the following applied. Wages were measured by Kroner per hour; they are represented by *actual* wages for working women and by *potential* wages for women who did not work at the time of the interview.<sup>5</sup> Experience was measured by the number of years worked since completing the highest education. Finally, education was measured by years of schooling.<sup>6</sup>

The results for both models are presented in table 1. Part (a) of the table contains the results before correcting for heteroskedasticity whereas part (b) shows the results after heteroskedasticity has been corrected for. The details of the correction are as follows. By reference to the large-sample procedure described in section III, we postulate that heteroskedasticity in the earnings equation takes the form

$$\sigma_i^2 = \alpha S_i^\delta \quad (21)$$

and obtain the following estimates:

$$\hat{\alpha} = \begin{matrix} 0.429, \\ (0.173) \end{matrix} \quad \hat{\delta} = \begin{matrix} -0.610 \\ (0.112) \end{matrix} \quad (F = 15.52*).$$

(\*Significant at the 1% level.) Correction for heteroskedasticity was then implemented in accordance with equation (15).

<sup>5</sup> Potential wage was measured by the answer given to the question “If you were to be working just now, how much would you be paid per hour?”

<sup>6</sup> When applying the LOF test to model 2 we also used dummy variables to represent various kinds of education instead of the linear and quadratic number of years of schooling. Since the results were very close for both formulations, we present only the latter.

TABLE 1.—REGRESSION RESULTS AND LACK-OF-FIT COMPUTATIONS FOR MODELS 1 AND 2

(a) Before Correcting for Heteroskedasticity								
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	SSE	SSP	F
Model 1 ( $n = 1958, m = 154, K = 4$ )	2.23 (0.05)	0.031 (0.006)	-0.00068 (0.00024)	0.078 (0.004)	—	283.2	255.7	1.29 <sup>b</sup>
Model 2 ( $n = 1958, m = 154, K = 5$ )	2.95 (0.16)	0.030 (0.006)	-0.00066 (0.00024)	-0.051 (0.028)	0.006 (0.001)	280.1	255.7	1.15
(b) After Correcting for Heteroskedasticity								
	$\hat{\beta}_0$	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	SSE*	SSP*	$\chi^2$
Model 1 ( $n = 1582, m = 68, K = 4$ )	2.34 (0.05)	0.042 (0.007)	-0.0013 (0.0004)	0.066 (0.003)	—	1 616.7	1 514	102.7 <sup>a</sup>
Model 2 ( $n = 1582, m = 68, K = 5$ )	3.26 (0.18)	0.036 (0.007)	-0.0010 (0.0004)	-0.104 (0.032)	0.0080 (0.0014)	1 588.8	1 514	74.8

Note: Standard errors are in parentheses.

<sup>a</sup> Significant at the 5% level.

<sup>b</sup> Significant at the 1% level.

The most important result of table 1 is that Model 1 is rejected by the LOF test whereas Model 2 passes the test. This holds whether a correction for heteroskedasticity is carried out or not. Thus the evidence of Mincer and Polachek (1974) is confirmed by our results. The linear form of schooling in the earnings equation for married women appears to be inappropriate according to our evidence.<sup>7</sup>

We conclude by giving a more detailed consideration to the marginal effect of schooling on the earnings of women. Using the results for Model 2 after correcting for heteroskedasticity, we estimate the marginal effect of schooling as

$$\left[ \frac{\partial E(\ln W_{ij})}{\partial S_i} \right]_{X_i \text{ constant}} = -0.104 + 0.016S_i. \quad (22)$$

This result is consistent with that of Mincer and Polachek (1974) in that the effect of schooling is an increasing function of schooling. In addition, the marginal effect of schooling is positive for  $S_i \geq 6.5$  years. As there are at least 8 years of compulsory primary school-

<sup>7</sup> It is, of course, possible that the term  $S_i^2$  captures the influence of other control variables such as those considered by Mincer and Polachek (1974). We do not pursue this further since our main purpose is to illustrate the use of the LOF test.

ing, these data suggest that the marginal effect of schooling (for fixed experience) is positive. This is consistent with the result which has been found for men.

#### REFERENCES

- Bartlett, M. S., and D. G. Kendall, "The Statistical Analysis of Variance Heterogeneity and the Logarithmic Transformation," *Journal of the Royal Statistical Society, Series B*, 8 (1946), 128-138.
- Battese, George E., "Agricultural Economists, Response Functions and Lack-of-fit tests," *Review of Marketing and Agricultural Economics* 45 (1977), 85-94.
- Behrman, Jere R., and Nancy Birdsall, "The Quality of Schooling: Quantity Alone Is Misleading," *American Economic Review* 73 (Dec. 1983), 928-946.
- Central Bureau of Statistics of Norway, 1981, Fertility survey 1977 (Oslo) (1981).
- Chiswick, Carmel U., "Analysis of Earnings from Household Enterprises: Methodology and Application to Thailand," this REVIEW 65 (Nov. 1983), 658-662.
- Griliches, Zvi, "Estimating the Returns to Schooling: Some Econometric Problems," *Econometrica* 45 (Jan. 1977), 1-22.
- Mincer, Jacob, *Schooling, Experience and Earnings* (New York: Columbia University Press, 1974).
- Mincer, Jacob, and Solomon Polachek, "Family Investments in Human Capital: Earnings of Women," *Journal of Political Economy* 82 (Mar./Apr. 1974), 76-108.
- Ryder, Harl E., Frank P. Stafford and Paula E. Stephan, "Labor, Leisure and Training over the Life Cycle," *International Economic Review* 17 (Oct. 1976), 651-674.