# MEASUREMENT ERROR IN SELF-REPORTED HEALTH VARIABLES

J. S. Butler, Richard V. Burkhauser, Jean M. Mitchell, and Theodore P. Pincus*

*Abstract*—Measurement error may be an important source of bias in studies using self-reported health indicators to explain work behavior. As a test of measurement error, the tetrachoric correlation coefficient is used to examine the relationship between two alternative measures of arthritis, a standard self-reported measure and a simulated clinical measure. While the two measures are highly correlated, measurement error is found. Regression analysis demonstrates that it varies systematically across different socioeconomic groups. In particular, individuals who are not working tend to report their health incorrectly, perhaps owing to social pressure to justify not having a job.

THE role of health in behavioral models is an important and controversial empirical question. Economists have modelled health as a determinant of the utilization of health services, as well as in work and retirement decisions. Almost all empirical research in these literatures has relied on some form of subjective health measure. Often this measure was a simple dichotomous variable, such as, "Does your health limit the kind or amount of work you can do?" Examples of papers using this type of indicator include Bartel and Taubman (1979), Wilensky and Cafferata (1983), Diamond and Hausman (1984), and Gustman and Steinmeyer (1986). A criticism of such measures of health is that they may bring systematic biases and reporting error into the analysis. Undoubtedly, since self-reported health reflects perceived health, it may measure something different from actual health. Still, considerable evidence suggests that self-reported measures are generally valid indicators of actual health. See Maddox and Douglas (1973), Ferraro (1980), and Waldron et al. (1982).

Researchers have adopted alternative methods to control for differences in health status when estimating models of labor supply. Da Vanzo et al. (1976), for instance, excludes unhealthy individuals from the analysis. With this approach one assumes that unhealthy persons are in a state of disequilibrium and as a consequence are off their long-run labor supply curves. An alternative somewhat more satisfying explanation is that unhealthy workers, those with chronic health problems, probably have permanently lower labor supply levels. Regardless of the justification used, since the remaining healthy individuals represent a selected group, the findings have only limited generalizability.

A second approach is to incorporate the self-reported measure of health as an exogenous variable in a labor supply equation. Diamond and Hausman (1984), for instance, followed this strategy in a longitudinal study of retirement behavior. Their results showed that a deterioration from good to poor health is the most significant factor determining when older workers will retire. A problem with this strategy is illustrated by the following case. Suppose that two individuals both report a health impairment which limits their ability to work. One of these individuals, however, has continued to work despite this health limitation, whereas the other has dropped out of the labor force. In this circumstance the health indicator is not controlling for variations in health status between the two individuals.

A third way researchers have acknowledged the role of health in models of work behavior is to partition the sample based on the response to a self-reported health question. Gustman and Steinmeyer (1986), for example, distinguish between healthy and unhealthy persons. They find that those reporting good health as well as those reporting poor health responded to the increased work incentives of the 1983 Social Security Amendments. Persons reporting good health, however, were much more responsive to these incentives. This treatment of health is troublesome because considerable unexplained variation linked to differences in health status exist within each group. To recognize such differences in health status, additional control variables for health should be incorporated into the regression equation for each group. In fact, some recent studies have adopted such a strategy when estimating either a labor supply or earnings function. See Lambrinos (1981), Chirikos and

Received for publication January 12, 1987. Revision accepted for publication April 30, 1987.

* Vanderbilt University, Vanderbilt University, Johns Hopkins University, and Vanderbilt University, respectively.

Nestel (1985), Johnson and Lambrinos (1985), and Mitchell and Butler (1986). Generally, the findings from these studies demonstrate that the additional controls for health have a significant influence.

Recent efforts to assess the importance of the bias caused by using self-reported measures of health in behavioral models involve estimation of an empirical model which incorporates health status as an independent variable. See Lambrinos (1981), Manning et al. (1982), Anderson and Burkhauser (1985), and Bazzoli (1985). Results using the self-reported measure of health are then compared with estimates employing more objective indicators of health. Since these objective measures often reflect a scale of disability or actual mortality over a subsequent period, they are assumed to approximate actual health more closely. These studies suggest that models using self-reported health yield significantly different results from those employing more comprehensive or objective measures of health.

This evidence illustrates the value of using more objective indicators of health. This methodology, however, presents two potential problems. First, it relies on the assumption that the more objective measure is a better indicator of actual health. It is possible, however, that both measures are poor indicators of health status.[1] Second, it does not explicitly examine the relationship between the two indicators. Thus the potential bias due to measurement error in the dichotomous self-reported measure is not assessed. The extent of this bias is of particular concern to researchers who are often forced to use self-reported health measures because of data limitations.

Here the importance of measurement error in a dichotomous self-reported health variable of the type frequently used in empirical models of work behavior is estimated. The 1978 Survey of Disability and Work is used to examine the relationship between a self-reported dichotomous measure of arthritis and a simulated clinical indicator based on symptoms of the disease. The tetrachoric correlation coefficient is estimated to assess the association between the two definitions of arthritis. Multiple regression analysis is then used to determine the degree to which the accuracy of reporting arthritis by survey respondents, as measured by the tetrachoric correlation coefficient, is linked to specific socioeconomic characteristics.

**Two Indicators of Arthritis**

The 1978 Survey of Disability and Work was conducted by the Social Security Administration. Information is available for a sample of 5,652 persons, representative of the U.S. civilian, noninstitutionalized population between the ages of 18 and 64 in 1978. Detailed information is available on the socioeconomic characteristics, labor force behavior and health status of these working age individuals. The detailed questions regarding health enable us to investigate the extent of measurement error in one specific health condition —arthritis.

Respondents were asked the following question pertaining to their health status: "Which of the following conditions or illnesses do you have now that a doctor has told you about?"[2] One of the 37 possible health conditions read by the interviewer was "Arthritis or Rheumatism." A response of "yes" or "no" was recorded for each subject. All individuals responding "yes" to this question were classified positive for a variable termed *ARTHRITIS DIAGNOSIS*.

The presence and the severity of arthritis can be assessed using an index of pain, swelling, and stiffness of specific joints, analogous to a clinical joint count. See Camp (1971) and McCarty (1979). In this procedure, the number of affected joints is enumerated. The disability survey includes questions regarding pain and swelling of specific joints. Respondents were asked: "Does pain in any part of your body bother you enough to be a problem?" Those who responded positively were asked: "Where do you have the pain?" Possible response options include the hip, knee, ankle, foot, shoulder, elbow, hand, neck and back. Respondents were asked similar questions regarding swelling and stiffness. This detailed information regarding pain and swelling of specific joints was used to construct an alternative measure of arthritis. Subjects who reported pain in the hip, knee, ankle, foot, shoulder, elbow, or hand, or swelling in one of

[1] See, for instance, the exchange between Haveman and Wolfe (1984) and Parsons (1984) regarding the problems of using subsequent mortality as a measure of current health.

[2] Economists looking at the effect of the disability transfer program on labor supply have used this self-reported question as a measure of health. See Haveman, Wolfe, and Warlick (1984), and Haveman, Wolfe, and Warlick (1986).

these joints were classified as having *ARTHRITIS SYMPTOMS*.[3] Information on stiffness was also included in the joint count when an individual reported pain or swelling in at least one joint, but excluded when only stiffness was reported.[4]

Although the variable *ARTHRITIS SYMPTOMS* is also based on self-reported information, we maintain that it is a better indicator of arthritis because it eliminates some of the sources of measurement error inherent in the dichotomous survey question. First, since the alternative definition simulates a clinical procedure which focuses on symptoms of the disease, it is a more objective assessment of arthritis. Second, this definition captures individuals who may actually have arthritis but not be aware of the condition because they rarely visit a doctor. Third, this measure may account for some individuals who do not report an arthritic condition because they consider the disease socially stigmatic. Finally, this definition may exclude some respondents who state they have arthritis as a justification for not working.[5]

### The Tetrachoric Correlation Coefficient

The correlation coefficient, the standard measure of association between two variables, is the covariance between the two, normalized by division by their standard errors. For the normal distribution, the correlation coefficient is a basic parameter. This is not the case for most distributions. The prevalence of the normal distribution in statistical theory explains the importance of the correlation coefficient.

The best-known correlation coefficient is the Pearson product-moment correlation which assumes the data represent the variables directly. If the data are actually dichotomized versions of underlying normally distributed variables, then the product-moment correlation is inappropriate. It amounts to regressing one poorly measured dichotomized variable on another. Measurement error causes the estimated correlation to be biased toward zero. Because data used to indicate health status in empirical studies are often dichotomous, they are subject to this type of bias.

The tetrachoric correlation coefficient is the maximum likelihood estimator of the correlation coefficient between two joint normally distributed random variables observed dichotomously. It avoids the "measurement error" which attenuates the product moment correlation. As a maximum likelihood estimator, it is asymptotically normally distributed with an estimable variance. We use the variance of the estimated tetrachoric correlation coefficient, which depends on the sample size and the distribution of the four cells of the table, in the subsequent regression analysis.[6] See Kendall

[3] We recognize that the variables *ARTHRITIS DIAGNOSIS* and *ARTHRITIS SYMPTOMS* include individuals who do not meet criteria for the medical definition of arthritis (i.e., inflammation of joints). The 1978 Survey of Disability and Work from which we derive our *ARTHRITIS DIAGNOSIS* variable used the broader health condition category "Arthritis or Rheumatism." Rheumatism is an imprecise term which encompasses a number of health conditions other than arthritis (e.g., bursitis, tendonitis). Nevertheless, these non-arthritic conditions are most likely also characterized by pain, swelling, and stiffness of joints and, therefore, will also be captured by the *ARTHRITIS SYMPTOMS* variable. Throughout the paper we use the word arthritis rather than arthritis and rheumatism for ease of exposition.

[4] One advantage of the measure *ARTHRITIS SYMPTOMS* is that persons with particular forms of arthritis can be identified and their incidence in our population compared with that in other studies. For instance, since subjects were reviewed for the presence of pain and swelling in both the right and left joints, we could distinguish those individuals with symptoms of rheumatoid arthritis. Persons reporting symmetrical pain or swelling in two or more specific joints, for example both hands and both knees, were classified as having a condition termed "Symmetrical Polyarthritis," a form of the disease which may be regarded as a surrogate for rheumatoid arthritis. The prevalence of Symmetrical Polyarthritis found in this representative sample of the working age population is 2.4% of females and 1.4% of males. See Mitchell, Burkhauser, and Pincus (in press). This finding is consistent with the prevalence rates for rheumatoid arthritis documented by previous studies in the epidemiology literature. See Mikkelsen et al. (1967), Camp (1971), Kelsey (1982), and Lawrence and Shulman (1984).

[5] The variable *ARTHRITIS SYMPTOMS* simulates an established diagnostic procedure. We have excluded from this definition joint problems which may be linked to diseases other than arthritis. For example, pain in the back and neck were excluded because these symptoms are frequently attributable to disorders other than arthritis. Similarly, swelling of the ankle was excluded because this condition may be characteristic of other conditions. Swelling of the hip was not included because it was not reported on the questionnaire.

[6] In judging the appropriateness of using the tetrachoric correlation coefficient versus, for instance, the better known Pearson correlation coefficient two additional issues must be considered. First the underlying scale may be skewed. That is, most individuals have no or only a few inflamed joints while a few individuals have many inflamed joints. In such a case, the underlying scales are not normally distributed. Second, because the observed data are dummy variables, the Pearson correlation coefficient appears to be a better measure of the relationship between the variables. We argue that it is plausible that dummy variables used to identify a disease like arthritis, which is in most cases not a discrete infection but inflammation or wear-and-tear of joints, must in fact represent an

and Stewart (1979) for more information, including a discussion of estimating the tetrachoric correlation coefficient.

This paper focuses on one specific health condition—arthritis—and uses the tetrachoric correlation coefficient to assess the extent of measurement error in self-reported data. Both indicators of arthritis are dichotomized and represent an underlying scale of severity. Some respondents report that a doctor informed them of their condition, and some are classified as having arthritis because of symptoms, but each represents an underlying scale.

The potential bias in empirical models caused by measurement error is an important statistical concern. The standard econometric analysis of errors in variables indicates that the resulting bias in the estimated coefficients of a regression increases as the ratio of the variance of the measurement error to the variance of the variable measured with error increases. See Goldberger (1964, pp. 282–284). The higher the correlation is between the true and the measured exogenous variables, the lower is the bias in all estimated coefficients. The bias could be eliminated if the correlation between the variables, or any equivalent information, such as the variance of the measurement error, is known.[7] The bias is complicated in the case of health because its measurement error may be correlated with several variables which are likely to appear in any empirical model.

### A Test for Measurement Error

The tetrachoric correlation coefficient is estimated to ascertain the relationship between the two indicators of arthritis across different socioeconomic groups. The sample consists of 4,867 whites of working age, 18–64.[8] Generalized least squares regression analysis is used to test the importance of particular socioeconomic characteristics in explaining reporting error. The tetrachoric correlations classified by their relevant socioeconomic characteristics are used as the dependent variable in the regression. The independent variables account for socioeconomic characteristics frequently examined in empirical studies of individual behavior. These controls include employment status, income, age, education, marital status, annual doctor visits and health characteristics. Each tetrachoric correlation coefficient is weighted by the inverse of its standard error to correct for heteroskedasticity.[9]

This regression analysis is not structural. We have no general theory of how magnitudes of measurement error arise, although in some cases we propose explanations for the estimated effects. Rather, we are interested in predicting the severity of the problem for different socioeconomic groupings which appear in studies of labor market behavior. The estimated correlations are used as the dependent variable in the regression rather than the implied $z$-scores. By predicting the correlations, we can assess the degree of measurement

underlying scale with unlimited correlations. Thus, the product moment correlation, which is bounded away from plus and minus one with threshold variables, is less appropriate. Although the underlying scale cannot be measured, one can, in principle, normalize it to obtain a normal distribution which allows the use of the tetrachoric. Thus, a concave function of the index is the basis of the model. Although the scale cannot be seen, such transformations to obtain normal distributions have a long history in statistics.

[7] In order to make this correction, it is necessary to know the sum of squares and cross-products of the regressions. Publication of their values together with regression coefficients would allow the bias caused by traditional self-reported measures of health to be estimated.

[8] If one of the cells of the classification is empty, the tetrachoric correlation must be estimated as plus or minus unity. In finite samples, there may be no one showing a particular combination. When both variables represent a scale, every range of values, such as those for which a respondent does not have arthritis but says a doctor diagnosed it, has a positive probability. To evaluate the degree of association between variables for various subgroups of the population, the tetrachoric correlation coefficient is calculated only for those subgroups which have observations in all four cells. The other groups are assumed to fail because of sampling and thus are excluded from the analysis. The inclusion of nonwhites would have entailed too many such cells to permit a meaningful analysis. Furthermore, although the product moment correlation can be calculated when this situation occurs, it will not solve the problem caused by missing data, because it is also biased.

[9] Any positive definite weighting matrix, including the identity, produces consistent regression estimates but not consistent variances. The cells are quite heterogeneous, which can affect the estimated variances of the regression coefficients considerably. Since the dependent variable is the tetrachoric correlation coefficient of a cell, we feel the best weight to use is its estimated standard error. Hence, the weight reflects the sample size and the distribution of the responses determining the tetrachoric correlation coefficient. It is not conditioned on the exogenous variables. But we assume that the weights are proportional to the variance of the disturbance in the regression.

error in the indicator *ARTHRITIS DIAGNOSIS*.[10]

## Empirical Results

The relationship between the two arthritis measures is reported in table 1. *ARTHRITIS DIAGNOSIS* is the most prevalent condition among the 37 possibilities listed in the questionnaire. More than 11% of the sample reported an arthritic condition diagnosed by a doctor. In comparison, approximately 10% of the sample reported *ARTHRITIS SYMPTOMS*, the indicator analogous to the clinical joint count.[11] Less than 11% of the sample were incorrect matches: 6% of those who reported *ARTHRITIS DIAGNOSIS* reported no *ARTHRITIS SYMPTOMS* while 5% reporting *ARTHRITIS SYMPTOMS* reported no *ARTHRITIS DIAGNOSIS*. While the two definitions are not perfectly matched, the results show a strong positive tetrachoric correlation of 0.703. The standard error of the estimate is 0.016.

The tetrachoric correlations were calculated for different subgroups classified by their relevant socioeconomic characteristics. A strong positive correlation exists for most subgroups. The estimated correlations range from −0.58 to 0.98. Only 17 of the 59 correlations have values below 0.50. Two are negative. These estimates are used as the dependent variable in our regression analysis and are available from the authors upon request. Table 2 presents the definition of the variables used in the regression. All variables

TABLE 1.—CLASSIFICATION OF THE SAMPLE BY THE TWO INDICATORS OF ARTHRITIS

| *ARTHRITIS DIAGNOSIS* | *ARTHRITIS SYMPTOMS* | Percentage of Sample | Weighted Cell Size |
|---|---|---|---|
| yes | yes | 5 | 5,698 |
| no | no | 84 | 92,633 |
| no | yes | 5 | 5,622 |
| yes | no | 6 | 6,813 |

Source: Developed from data reported in the 1978 Survey of Disability and Work.

Note: Tetrachoric correlation coefficient = 0.703, Standard error of the estimate = 0.016. Total Sample = 4,867. The weighted cell size refers to the population represented by the individuals sampled and is based on sampling weights in the survey.

TABLE 2.—DEFINITION OF VARIABLES

| | |
|---|---|
| Work | Binary variable equals one if respondent is working. |
| High School Ed. | Binary variable equals one if respondent has 12 years of schooling. |
| College Ed. | Binary variable equals one if respondent has more than 12 years of schooling. |
| Income[a] | Binary variable equals one if family income is greater than $15,000. |
| Age (45–64) | Binary variable equals one if respondent's age is between 45 and 64. |
| Female | Binary variable equals one if respondent is female. |
| Married | Binary variable equals one if respondent is married with spouse present. |
| Joint Count[b] | The mean number of affected joints in a subgroup. |
| No Symptoms | Percentage of a subgroup who do not have arthritis according to the clinical indicator. |
| Doctor Visits | The mean number of annual doctor visits in a subgroup. |
| Constant | Subgroup of men aged 18–44 with less than 12 years of schooling, not married, not working, with income below $15,000. |

[a] Income includes money from jobs, net income from business or farm, pensions, dividends, interest, net rent, Social Security payments, and any other money income received by members of this family who are 16 years of age or older. The median total family income reported by members of the working age population in 1978 was between $15,000 and $19,999. We selected $15,000 as the benchmark to divide the sample by reported family income level

[b] Joint Count measures the mean number of affected joints of those individuals in each of the 59 subgroups who are diagnosed as having *ARTHRITIS SYMPTOMS*. One point was assigned for each joint reported to be painful in the following places: hip (2), knee (2), ankle (2), shoulder (2), elbow (2), hand (2), foot (2) and neck (1). Similar counts were recorded for swelling. One point was assigned for each separate joint reported to be swollen in the following places: knee (2), hand (2), shoulder (2), elbow (2), foot (2). If an individual reported pain or swelling in at least one joint we added 13 other possibilities for stiffness. These additional joints include stiffness in the following places: hip (2), knee (2), ankle (2), shoulder (2), elbow (2), hand (2), and neck (1).

[10] The tetrachoric correlation approach is an alternative to log-linear analysis and for our purposes is equivalent to an "inter-rater reliability" analysis. A log-linear analysis can be used whenever the probabilities of the cells of a table are to be explained (Bishop, Fienberg, and Holland, 1975). The logarithm of the probability is expressed as a linear function of explanatory variables. Our goal, however, is to estimate the correlation coefficient between the variables defining the table in order to assess the degree of measurement error caused by using an inexact indicator. We are not interested in, for example, the probability of having a large number of inflamed joints and having an arthritic condition which has been diagnosed by a physician. Log-linear analysis is not intended to estimate correlation coefficients. The model in an inter-rater reliability or closely related factor-analytic or MIMIC model (see Joreskog and Goldberger, 1975) is essentially the same as ours, but our approach focuses on a different aspect of the problem. A MIMIC model assumes that two or more indicators of an underlying true factor are observed. The indicators are used to derive the true factor. Psychologists frequently employ this methodology, e.g., to assign intelligence levels to individuals where the indicators are test scores.

[11] The population estimates for *ARTHRITIS DIAGNOSIS* and *ARTHRITIS SYMPTOMS* are consistent with previous figures found in epidemiology studies. See Mikkelsen et al. (1967), Camp (1971), Kelsey (1982), and Lawrence and Shulman (1984).

TABLE 3.—GENERALIZED LEAST SQUARES REGRESSION RESULTS OF SELF-REPORT RELIABILITY MODEL (SAMPLE SIZE = 59)

| Independent Variables | Coefficient (standard errors in parentheses) |
|---|---|
| Constant | 0.401 (0.484) |
| Work | 0.276[a] (0.099) |
| High School Ed. | 0.240[a] (0.061) |
| College Ed. | 0.097 (0.070) |
| Income | 0.125[a] (0.046) |
| Age (45–64) | 0.161 (0.084) |
| Female | −0.117 (0.060) |
| Married | 0.019 (0.062) |
| No Symptoms | −0.372 (0.543) |
| Joint Count | 0.038[b] (0.017) |
| Doctor Visits | 0.043 (0.142) |

Note: $\bar{R}^2 = 0.95$; $F = 95.62$.
[a] Significant at the 1% level.
[b] Significant at the 5% level.

except the indicator of arthritis severity and the frequency of doctor visits are discrete dichotomous variables.

The results of the generalized least squares regression are presented in table 3. The control variables in this model account for most of the variation in the measure of reliability. Those individuals who are working are more likely to report their health correctly than are respondents who are not working. The work variable is significant at the 1% level. Most other variables are insignificant at the 5% level or above. Note especially that the mean number of annual doctor visits for each subgroup does not significantly affect reporting accuracy. The joint count, however, is significant at the 5% level. Those who report a greater number of joints which are painful, swollen, or stiff are more likely to accurately report arthritis.

Our finding concerning the effect of employment status on reporting behavior is consistent with the view that individuals who are not working tend erroneously to report a disease or to classify their health as poor, perhaps as a justification for not having a job. The fact that the traditional self-reported health measure is less consistently related to the measure which simulates a clinical procedure for those not working must be of concern to researchers who use traditional self-reported health variables in behavioral models of work effort. The role of health may be exaggerated in the response of non-working individuals.

## Conclusion

Self-reported measures of health are routinely used in behavioral models. Their validity has been challenged because they reflect perceived health which may differ considerably from actual health. Using data developed from the 1978 Survey of Disability and Work we gauge the importance of measurement error in a traditional self-reported indicator of a chronic disease, arthritis. We estimate a tetrachoric correlation coefficient to ascertain the association between a traditional self-reported measure of arthritis and a simulated clinical measure of the disease. This alternative measure of arthritis, while derived from self-reported information, is superior because it focuses on symptoms of arthritis—namely, pain, swelling or stiffness of specific joints. It is hence analogous to a clinical procedure employed by rheumatologists to diagnose the presence and severity of arthritis. The tetrachoric correlation coefficient estimates the underlying correlation between the two indicators of this disease.

The two measures of arthritis show a strong positive relationship. In general, a strong correlation is also observed when we classify the sample by socioeconomic characteristics. This suggests that in general both measures identify those who have the disease. In this respect our results are consistent with others who have found traditional self-reported measures to be generally valid indicators of actual health. Nonetheless, traditional self-reported measures must be used with caution. The regression results demonstrate that work status significantly affects measurement error. Individuals who are not working are more likely to err when self-reporting the presence or absence of an arthritic condition. Thus, measurement error in such self-reported health variables may represent a source of significant bias in models used to explain work behavior.

REFERENCES

Anderson, Kathryn H., and Richard V. Burkhauser, "The Retirement-Health Nexus: A New Measure of an Old Puzzle," *Journal of Human Resources* 20 (Summer 1985), 315–330.

Bartel, Ann, and Paul Taubman, "Health and Labor Market Success: The Role of Various Diseases," this REVIEW 61 (Feb. 1979), 1–8.

Bazzoli, Gloria J., "The Early Retirement Decision: New Empirical Evidence on the Influence of Health," *Journal of Human Resources* 20 (Spring 1985), 214–234.

Bishop, Yvonne M., Steven E. Fienberg, and Paul W. Holland, *Discrete Multivariate Analysis: Theory and Practice* (Cambridge, MA: The MIT Press, 1975).

Camp, A. V., "An Articular Index for the Assessment of Rheumatoid Arthritis," *Orthopaedics* 4 (1) (1971), 39–45.

Chirikos, Thomas N., and Gilbert Nestel, "Further Evidence on the Economic Effects of Poor Health," this REVIEW 67 (Feb. 1985), 61–69.

Da Vanzo, Julie, Dennis De Tray, and David H. Greenberg, "The Sensitivity of Male Labor Supply Estimates to Choice of Assumptions," this REVIEW 58 (Aug. 1976), 313–325.

Diamond, Peter A., and Jerry A. Hausman, "The Retirement and Unemployment Behavior of Older Men," in H. Aaron and G. Burtless (eds.), *Retirement and Economic Behavior* (Washington, D.C.: The Brookings Institution, 1984).

Ferraro, Kenneth F., "Self-Ratings of Health Among the Old and Old-Old," *Journal of Health and Social Behavior* 20 (1980), 377–383.

Goldberger, Arthur S., *Econometric Theory* (New York: John Wiley & Sons, Inc., 1964).

Gustman, Alan L., and Thomas L. Steinmeyer, "A Disaggregated, Structural Analysis of Retirement by Race, Difficulty of Work and Health," this REVIEW 68 (Nov. 1986), 509–513.

Haveman, Robert, and Barbara Wolfe, "The Decline in Male Labor Force Participation: Comment," *Journal of Political Economy* 92 (June 1984), 532–542.

Haveman, Robert, Barbara Wolfe, and Jennifer Warlick, "Disability Transfers, Early Retirement and Retrenchment," in H. Aaron and G. Burtless (eds.), *Retirement and Economic Behavior* (Washington, D.C.: Brookings Institution, 1984).

______, "Behavior Responses to Social Security Retrenchment: Estimates from a Trichotomous Choice Model," Institute for Research on Poverty DP 789-85 (1986).

Johnson, William G., and James Lambrinos, "Wage Discrimination Against Handicapped Men and Women," *Journal of Human Resources* 20 (Spring 1985), 264–277.

Joreskog, Karl, and Arthur S. Goldberger, "Estimation of a Model with Multiple Indicators and Multiple Causes of a Single Latent Variable Model," *Journal of the American Statistical Association* 70 (1975), 631–639.

Kelsey, Jennifer L., *Epidemiology of Musculoskeletal Disorders*, *Monographs in Epidemiology and Biostatistics*, Vol. 3 (New York: Oxford University Press, 1982).

Kendall, Maurice, and Alan Stuart, *The Advanced Theory of Statistics*, Vol. 2 (New York: Macmillan Inc., 1979).

Lambrinos, James, "Health: A Source of Bias in Labor Supply Models," this REVIEW 63 (May 1981), 81–98.

Lawrence, Reva C., and Lawrence E. Shulman, *Current Topics in Rheumatology: Epidemiology of the Rheumatic Diseases* (New York: Gower Publishing, 1984).

Maddox, George L., and Elizabeth B. Douglas, "Self-Assessment of Health: A Longitudinal Study of Elderly Subjects," *Journal of Health and Social Behavior* 14 (1973), 87–93.

Manning, William G., Joseph P. Newhouse, and John E. Ware, "The Status of Health in Demand Estimation: Or Beyond Excellent, Good, Fair or Poor," in V. Fuchs (ed.), *Economic Aspects of Health* (Chicago: University of Chicago Press, 1982).

McCarty, Daniel J., *Arthritis and Allied Conditions* (Philadelphia: Leas and Feibiger, 1979).

Mikkelsen, William M., Horace J. Dodge, Ivan F. Duff, and Hiroo Kato, "Estimates of the Prevalence of Rheumatic Disease in the Population, Tecumseh, Michigan 1959–1960," *Journal of Chronic Disease* 20 (1967), 351–369.

Mitchell, Jean M., Richard V. Burkhauser, and Theodore P. Pincus, "The Importance of Age, Education, and Co-Morbidity in the Earnings Losses of Individuals with Symmetrical Polyarthritis," *Arthritis and Rheumatism* (in press).

Mitchell, Jean M., and J. S. Butler, "Arthritis and the Earnings of Men: An Analysis Incorporating Selection Bias," *Journal of Health Economics* 5 (Mar. 1986), 81–98.

Parsons, Donald O., "Disability Insurance and Male Labor Force Participation: A Response to Haveman and Wolfe," *Journal of Political Economy* 92 (June 1984), 542–549.

Waldron, Ingrid, Joan Harold, and David Dunn, "How Valid Are Self-Reported Measures for Evaluating the Relationship between Women's Health and Labor Force Participation?" *Women and Health* 7 (1982), 53–66.

Wilensky, Gail R., and Gail L. Cafferata, "Women and the Use of Health Services," *American Economic Review* 73 (May 1983), 128–133.