Soc. Sci. Med. Vol. 33, No. 2, pp. 127–137, 1991
Printed in Great Britain. All rights reserved

0277-9536/91 $3.00 + 0.00

EMPLOYEE AND JOB ATTRIBUTES AS PREDICTORS OF ABSENTEEISM IN A NATIONAL SAMPLE OF WORKERS: THE IMPORTANCE OF HEALTH AND DANGEROUS WORKING CONDITIONS

J. PAUL LEIGH

Department of Economics, San Jose State University, San Jose, CA 95192-0114, U.S.A. and Stanford Department of Medicine, Division of Immunology, 750 Welch Road, Suite 305, Palo Alto, CA 94304, U.S.A.

Abstract—This study reports on research which looks for employee and job characteristics which correlate with absenteeism. A large cross-sectional national probability sample of workers employed for at least 20 hr per week is analyzed ($n = 1308$). The dependent variable is the number of self-reported absences during the past 14 days. Thirty-seven independent variables are considered. Ordinary Least Squares (multiple regressions), two-limit Tobits, and two-part models are used to assess the statistical and practical significance of possible covariates. Statistically significant predictors included health variables such as being overweight, complaining of insomnia, and hazardous working conditions; job characteristics such as inflexible house; and personal variables such as being a mother with small children. Variables reflecting dangerous working conditions appear to be the strongest correlates of absenteeism. Notable variables which do not predict absenteeism include age, race, wages, and job satisfaction. Future research should direct attention toward workers' health and working conditions as covariates of absenteeism, since they are strongly significant in this study and have been neglected by most absenteeism investigators.

Key words—absenteeism, health, working conditions, Tobit

Studies on absenteeism by psychologists, occupational health researchers and economists are legion. Comprehensive reviews of the earlier psychological studies, which frequently focused on job satisfaction, are available [1, 2]. The most recent psychological studies have been concerned with measurement problems [3], statistical modeling issues involving censoring of the dependent variable [4, 5], causal modeling using Lisrel estimates [6], linking absenteeism to being tardy [7], investigating gender differences [8], and concerned with using psychiatric traits as predictors [9]. Some of the more recent studies are reviewed in Steers and Rhodes [2] and Farrell and Stamm [10]. The occupational health studies, by contrast, have tended to focus on environmental exposures [11], smoking [12], and workers' health [13]. A natural recent trend in absenteeism studies has been to link the two literatures. Parks [14], for example, reviewed the burgeoning occupational health and psychological literature on smoking. She also recently pulled together related literature on relative weight of the employee as a possible predictor [15].

A few studies have been conducted by economists who have been concerned primarily with economic incentives, unionization, and demographic traits as possible predictors and with statistical modeling issues. Allen [16, 17] and Winkler [18], for example, found that wages and paid sick leave were inversely related to absenteeism. Leigh [19] found evidence that union members were more likely to be absent than nonunion members, and unemployment to be positively associated with absenteeism. Paringer [20] found that gender differences in absenteeism were largely the result of a mother's tendency to stay home with her small children if they became ill. Finally, Leigh [21] noted, as Hammer and Landau [4] and Baba [5] did in prior studies, that the dependent variable, absence, is censored at zero, and that the proper estimating technique would be a single-limit Tobit regression rather than Ordinary Least Squares or multiple regression (MR). MR produces statistically inconsistent estimates when dependent variables are censored.

The economic studies have not been widely cited by psychologists or occupational medicine investigators, perhaps due to their inadequacies, but also perhaps due to the tendency of investigators in any field to read familiar journals. But, a multi-disciplinary approach can sometimes produce insights as Parkes [14, 15] has demonstrated. This study attempts to pull together a variety of hypotheses from psychology, occupational medicine and economics. The hypotheses are tested on a large national probability sample of full-time workers in the University of Michigan's *Quality of Employment Survey* for 1973 (QES). Thirty-seven possible covariates compete for explanatory power in predicting absenteeism. The approach is empirical and eclectic. No variable or set of variables receives special consideration. The sub-title of this paper was added only after the analyses were completed. Three of the favorite economic variables—wages, other income, and unemployment—for example, are not found to be important predictors. The surprisingly strong showing for worker health and hazardous working conditions variables was not anticipated.

Two recently developed econometric models, the two-limit Tobit and the two-part model, are used to search for covariates. An additional methodological contribution of the paper involves the method for assessing the importance of an independent variable. The statistical and practical significance of a variable will be used to judge that variable's importance [22]. Stepwise regression methods for assessing importance are eschewed [23].

DATA

In 1974, the University of Michigan's Survey Research Center administered its Quality of Employment Survey (QES) to a cross-section of 1445 employees. The QES is a U.S. national probability sample of employees who worked for pay for at least 20 hr per week. Kallenberg [24] provides a recent description of the sampling design and an outline of many of the studies that have relied on the QES. The data are old, but the quality is high and the extensive set of variables pertaining to job conditions is unsurpassed for a national probability sample. For example, recent studies by Cook [25] and Fennell *et al.* [26] looked at wages and alcohol consumption among QES workers. Robinson [27] used the QES to look for correlations between desires for unionization and job hazards. Despite that the QES was funded by the National Institute for Occupational Safety and Health, a recent review of published studies relying on QES data revealed far more use by sociologists and economists than occupational medicine investigators or industrial psychologists [24].

Not all of the 1445 respondents interviewed by the QES researchers are included in the analysis. Self-employed persons are excluded, following Allen [16, 17] and Leigh [19, 20]. Absenteeism carries less meaning for persons who are self-employed. The relevant sample size is thus reduced to 1308. Because of the high quality of the QES, very few missing data were encountered. But, in the rare cases of missing data, mean values are substituted.

Table 1 presents acronyms, definitions and means on each of the variables in the analysis. The first dependent variable, ABSENT14, was the number of

Table 1. Definitions of variables and means

		Mean
DEPENDENT VARIABLES		
ABSENT14	Number of days absent in the past 14 days, excluding vacation	0.465
BINARYABS	Equals 1 for any positive absence days and 0 otherwise	0.170
INDEPENDENT VARIABLES		
Demographic variables		
AGE	Age in years	39
AGE-SQUARED	Age-squared	1521
ED	Years of schooling	12.27
MARRIED	Equals 1 if married and living with spouse, 0 otherwise	0.748
NONWHITE	Equals 1 for non-whites, 0 otherwise	0.124
FEMALE	Equals 1 for females	0.321
KIDS < 6	Number of children under 6 years old living in the household	0.401
FEMALE*KIDS < 6	Interaction term of female with number of young kids	0.082
Health variables		
COLDFLU	Equals 1 if respondent reported one or more colds or flu in past 12 months, 0 otherwise	0.695
SLEEP	Equals 1 if respondent reported problems falling or staying asleep at night, 0 otherwise	0.457
FAT	Equals 1 for respondent whom the interviewer felt was overweight, 0 otherwise	0.197
DRINK	Average number of alcoholic drinks consumed in a year	115
SMOKE	Equals 1 if respondent smokes either cigarette, pipe or cigar, 0 otherwise	0.497
POORAPT	Equals 1 if respondent says has poor appetite	0.109
DISABLE	Equals 1 if respondent reports some work-limiting disability, 0 otherwise	0.092
BACKPAIN	Equals 1 if respondent reported frequent backpain	0.198
DANGER	Equals 1 if respondent reported being exposed to dangerous or hazardous conditions at work, 0 otherwise	0.434
INJURY	Equals 1 if respondent believes injury or illness was caused or aggravated by job	0.188
OCCDTH	Estimated probability of death within occupation × 100	0.052
Aspects of the job		
OVERTM	Equals 1 if there is frequent overtime	0.266
SAMEHRS	Equals 1 if works same hours, no flexibility with work hours	0.808
HOURS	Average weekly hours at work	40.9
VERYSAT	Equals 1 if respondent was 'very satisfied' with job	0.220
ORGSIZE	Number of employees at respondent's work site	494
REPETIT	Equals 1 if respondent reported job is repetitious, 0 otherwise	0.529
PHYEFF	Equals 1 if job requires 'lots of' physical effort, 0 otherwise	0.262
MONTHS	Months of tenure at present job	89
Economic incentives		
UNION	Equals 1 for union member, 0 otherwise	0.301
UNFAIRPR	Equals 1 if respondent feels promotions are unfair	0.153
PAIDVAC	Equals 1 if respondent reported job allows paid personal vacation, 0 otherwise	0.624
MATERPY	Equals 1 if job offers maternity pay	0.013
COMMUTE	Average number of minutes spent commuting to main job	22.46
MOONLITE	Equals 1 if respondent holds two jobs	0.124
PAIDSKLV	Equals 1 if job has paid sick leave	0.558
SHORTAGE	Equals 1 if respondent states there is a shortage of workers with his/her skills in immediate area	0.474
WAGE	Annual earnings of respondent	9621
OTHERINC	Other annual income received by family member of respondent, but not earnings from respondent's job. This variable includes other family members' wages, as well as rent, dividends, interest, but not transfer payment. OTHERINC is measured in thousands of dollars	3642

Source: The data tape (QES73) was made available by the Inter-University Consortium on Political and Social Research.

days during the past 14 that the respondent did not report to work. Of the 1308, 83% or 1066 did not report any absences. These individuals received a 0 for ABSENT14. There were 109 of the respondents who reported 1 absence; 51 who reported 2; 17 reported 3; 10 reported 4; 12 reported 5; 4 reported 6; 4 reported 7; 1 reported 8; 6 reported 10; 1 reported 12; and 6 reported 14. The first dependent variable is censored. ABSENT14 cannot be less than zero as a matter of logic, or greater than 14 because of design of the question. ABSENT14 is not normally distributed.

The second dependent variable is a binary 0–1 variable, BINARYABS, which equals "1" for persons with *any* absences and "0" otherwise. BINARYABS = 1 when ABSENT14 > 0, and BINARYABS = 0 when ABSENT14 = 0.

Four separate sets of possible predictors are considered. The first set includes demographic variables such as (1) age, (2) age-squared, (3) education, (4) marital status, (5) race, (6) gender, (7) number of small children and (8) and interaction variable of gender with children. The second set includes health and hazardous working conditions variables reflecting the respondent's (9) recent history of colds or flu, (10) sleep trouble, (11) relative weight, (12) alcohol use, (13) smoking status, (14) appetite, (15) disability status, (16) back pain, (17) subjective evaluation of how dangerous the job is, (18) subjective evaluation of influence of job on injuries and illnesses, and (19) probability of dying on the job. The third set involves aspects of the job and includes such variables as (20) frequent overtime, (21) inflexible hours, (22) average weekly hours, (23) job satisfaction, (24) organization size, (25) subjective evaluations of the repetitive nature of the job, (26) subjective evaluation of the physical effort involved with the job, and (27) months of tenure. Economic incentives comprise the fourth category and includes (28) union status, (29) subjective evaluation of the promotion process, (30) presence of paid vacation, (31) presence of maternity pay, (32) average commute, (33) whether respondent holds a second job, (34) sick leave, (35) subjective evaluation of demand for respondent's skill in geographic area, (36) annual earnings, and (37) other family income.

With the exception of OCCDTH, the other variables in Table 1 are self-explanatory. OCCDTH represents estimates of the probability of a job-related death within the respondent's 3-digit U.S. Census Occupation Code classification [28]. The annual chances of dying on-the-job typically range from roughly 1/10,000 to 1/1,000,000. Loggers, airplane pilots and asbestos workers have among the most dangerous jobs, while many desk jobs have mortality numbers close to zero.

METHOD

Ordinary least squares or multiple regression (MR) assumes that the dependent variable is continuous, normally distributed, and can range (in theory) from negative to positive infinity. ABSENT14 is not normal and cannot lie outside 0 and 14. Moreover, 83% of ABSENT14 observations are at zero. It can be shown that for variables such as ABSENT14, MR will produce parameter (slope) estimators that are statistically biased, even in large samples [29]. The Tobit model will produce unbiased estimates under these circumstances.

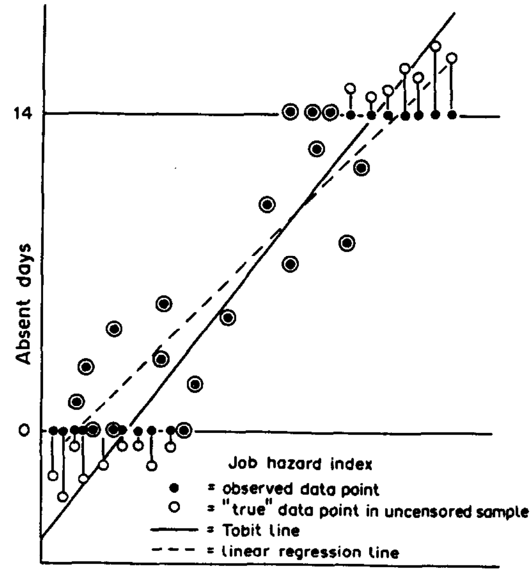

Fig. 1. Tobit vs linear regression.

Figure 1 illustrates the problem. Absenteeism is measured along the vertical axis, and an index of job hazards along the horizontal axis. Absenteeism is censored at 0 and 14. The Tobit model assumes that negative absenteeism and absence days beyond 14 are possible, but are not available in the data. The QES question did not allow for more than 14 days, hence the questionnaire resulted in censorship beyond 14. Negative absenteeism might correspond to carrying work home, arriving early, working on the weekends, working through lunch, working overtime, taking fewer breaks, and so on. Some people in the QES might, in fact, work through lunch and, thus, have negative absenteeism. Data points for these people are indicated with "Os" below the horizontal axis. The QES records these people with absence at zero, i.e. with one of the dots ("●") along the horizontal axis. A similar phenomenon occurs beyond 14. In both cases, the distance from the circle "O" to the dot "●" is the extent of measurement error. ABSENT14 is *censored* at 0 and 14. True values of independent variables, such as age and job hazards, are available from respondents at the limiting 0 and 14 values, but true values of the dependent variable beyond 0 and 14 are not available. If true values on the independent variables were not available, the sample would be *truncated*.

Ordinary least squares or multiple regression assumes no measurement error (no censoring) and yields biased estimates of the true slope and intercept. Tobit allows for the possibility of measurement error (censoring) and will produce unbiased parameter estimates. Mathematical details are available in Maddala [29]. While Fig. 1 suggests that linear regression will underestimate the true slope, in general, the extent of an under or overestimate cannot be determined *a priori*. Nevertheless, most empirical

OLS estimates are smaller in absolute value than most Tobit estimates [30].

Tobit models are very similar to Cox models developed by biostatisticians to analyze survival data in which years of life is truncated [31]. Survival data is truncated since true values for both dependent and independent variables are not available beyond the limit.

The Tobit was originally designed to handle variables censored at zero and with a long right tail. Hammer and Landau [4] appear to be the first to suggest the need for the Tobit in absenteeism research. But, ABSENT14 is also limited at 14. Thus, the two-limit Tobit appears most appropriate. The two-limit Tobit has been widely used by economists. Vijverberg [32], for example, investigated workers' allocation of work hours between working for somebody else vs self-employment. The hours of employment in the Vijverberg example are censored at 0 and 70 per week.

The two-limit Tobit procedure, available in the LIMDEP software package [33], is used to search for covariates of absenteeism.

While the two-limit Tobit has clear statistical advantages, the parameter estimates require a little explanation. The estimated coefficients are partial derivatives for the underlying uncensored, sometimes unobserved data. A partial derivative represents how a 1-unit change in an independent variable will be associated with an estimated change in the dependent variable holding other variables constant. For example, suppose that an estimated coefficient on years of schooling were −0.2. This would suggest that a 1-unit increase in schooling from, say, 11 to 12 years would be associated with a 0.2 reduction in uncensored absenteeism holding other variables constant. Partial derivatives are useful in assessing the importance of an independent variable in any regression model [22]. The Tobit coefficient estimates for the underlying uncensored relationship can be transformed to apply to observed data with a simple transformation. Greene [30] describes this transformation and shows that the partial derivative in a two-limit Tobit can be expressed as follows:

$$\frac{\partial \text{ ABSENT14}}{\partial \text{ ED}} = (\text{Prob}\,[L < \text{ABSENT14*} < U]) \cdot \beta_{\text{ED}}$$

where ABSENT14 is days absent; ED is years of schooling; β_{ED} is the estimated two-limit Tobit coefficient for ED; L is the lower limit, 0, and U is the upper limit, 14; 'Prob' indicates probability; and ABSENT14* indicates the values of ABSENT14 between L and U. The 'Prob' expression surrounded by parentheses can be estimated by the percent of observations between 0 and 14. This percent between the lower and upper limits is 17% in the QES. Thus, estimated partial derivative for the observed data is simply $0.17 \cdot \beta_{\text{ED}}$.

The answer to which is the more appropriate estimated partial derivative depends on the intended use. If researchers and business executives were concerned solely with observed absences between the limits, the transformed partial derivatives would be appropriate. If, on the other hand, attention was on absenteeism which could be negative when employees worked through lunch or possibly exceed 14 days, then the *un*transformed Tobit estimate would be the most appropriate. The *un*transformed coefficients are likely to be the most relevant to a business since *un*censored absenteeism would be a better measure of 'true' work effort. In the analyses that follow, the latter approach for *un*transformed estimates is adopted.

Tobit has critics, however. Hay, Len and Rohrer [34] suggest that two-part models should be competitors for Tobits. In the first part of the two-part model, the dependent variable is treated as a dummy, equalling "1" for everybody with, in our case, a positive number of absences, and equalling zero otherwise (BINARYABS). The first equation includes all respondents, and is estimated with Probit. Estimated partial derivatives for Probit are explained in Maddala [35]. The second equation is restricted to respondents with positive absences and estimated with Ordinary Least Squares (OLS) or MR. The dependent variable in the second equation, or second part of the two-part model, is ABSENT14. Hay, Len and Rohrer demonstrated that two-part models, while perhaps less theoretically desirable, can nevertheless outperform Tobits in terms of mean squared error.

In the analyses to follow, all three techniques will be attempted: (1) Ordinary Least Squares or MR for ABSENT14; (2) Two-limit Tobit for ABSENT14; and (3) a Two-part model—Probit for BINARYABS and Ordinary Least Squares or MR for ABSENT14.

In the initial analyses, all independent variables are included in every regression. In subsequent analyses, sets of variables, such as the health and hazardous working conditions variables or the job-related variables, are included and excluded at different times so that some assessment can be made of their relative importance. Residual sums of squares are compared between long regressions which include all four sets of variables, and short regressions which omit one group of variables. F-Statistics are calculated using the ratios of residual sums of squares from long and short regressions [35]. OLS and F-tests, rather than Tobit and Chi-squares, are used to assess the importance of the four groups since the results in columns 1 and 2 of Table 2 are so similar.

Asymptotic and finite-sample t-statistics are calculated as the ratio of the estimated Tobit or MR or Probit coefficients to their estimated standard errors. The t-statistics indicate whether a particular independent variable is statistically significant. The larger the t, the lower the probability of a type 1 error and the greater the chance of achieving statistical significance. The critical 0.05 values for ts in a large sample, such as the QES, are 1.645 for a one-tailed test and 1.96 for a two-tailed test. The decision to use a one- vs two-tailed test depends upon prior expectations for the variable. For example, it is widely believed that health problems should contribute to absenteeism. Thus, all of the health variables should be judged against the 1.645 standard. But, a variable such as ORGSIZE should be judged against the 1.96 standard because there are conflicting hypotheses about size of the firm or organization on absenteeism. On

Table 2. Ordinary least squares (multiple regression), Tobit, and two-part model estimates explaining absenteeism

	Estimated coefficient and absolute *t*-statistic			
	1	2	3	
			Two-part model	
			3a	3b
	OLS, dependent variable is ABSENT14	Two-limit Tobit, dependent variable is ABSENT14	Probit, dependent variable is BINARYABS	OLS, dependent variable is ABSENT14
Independent variables				
Constant	−0.210	0.060	0.189	−0.044
	(0.320)	(0.105)	(1.125)	(0.015)
Demographic				
AGE	0.148	0.100	−0.006	0.304*
	(0.951)	(0.760)	(0.937)	(2.396)
AGE-SQUARED	−0.0012	0.001	0.00007	−0.0036*
	(0.898)	(0.811)	(0.833)	(2.314)
EDUCATION	−0.036	−0.047	0.001	−0.071
	(0.259)	(0.473)	(0.241)	(0.844)
MARRIED	0.927	−0.888	−0.097	−3.541*
	(0.698)	(0.824)	(0.192)	(2.433)
NONWHITE	1.642	1.418	0.075	1.070
	(1.318)	(1.548)	(1.282)	(1.557)
FEMALE	−1.029	−1.035	0.008	−1.244
	(1.025)	(0.864)	(0.182)	(1.692)
KIDS < 5	−0.128	−0.188	0.027	−0.407
	(0.152)	(0.438)	(1.293)	(1.405)
FEMALE*KIDS < 5	2.419**	3.376**	0.083*	1.126**
	(2.927)	(3.732)	(1.982)	(2.883)
Health				
COLDFLU	0.014	−0.0059	0.067**	−1.263**
	(0.070)	(0.076)	(2.403)	(2.581)
SLEEP	1.417*	1.453*	0.009	1.019**
	(1.982)	(2.336)	(0.284)	(2.465)
FAT	1.329*	1.070*	0.046*	0.241
	(1.674)	(1.681)	(1.951)	(0.488)
DRINK	0.010	−0.0006	0.00001	0.0005
	(0.737)	(0.573)	(0.159)	(0.538)
SMOKE	−0.319	0.282	0.055**	−0.105
	(1.137)	(0.528)	(2.049)	(0.241)
POORAPT	−0.288	−0.223	0.008	0.183
	(0.722)	(0.278)	(0.199)	(0.307)
DISABLE	0.641	0.959	0.012	0.394
	(0.513)	(1.402)	(0.263)	(0.599)
BACKPAIN	0.540	0.452	−0.011	0.948*
	(1.199)	(0.682)	(0.348)	(1.976)
DANGER	1.377**	1.594**	0.058*	1.128**
	(3.327)	(2.850)	(2.082)	(2.621)
INJURY	1.106*	1.523*	0.081**	−0.0848
	(2.234)	(2.259)	(2.405)	(0.191)
OCCDTH	3.259*	4.941*	0.113*	1.250
	(1.756)	(1.891)	(1.726)	(0.942)
Job aspects				
OVERTM	0.163	0.900	0.029	0.441
	(0.749)	(1.515)	(1.015)	(0.937)
SAMEHRS	1.292	1.370*	0.064*	−0.583
	(1.421)	(2.028)	(2.057)	(0.883)
HOURS	−0.008	−0.0023	−0.002	−0.003
	(0.283)	(0.085)	(1.418)	(0.104)
VERYSAT	1.094	1.647	−0.017	1.575
	(0.519)	(0.311)	(0.190)	(1.372)
ORGSIZE	0.0009	0.0005	0.00001	0.0001
	(1.619)	(1.753)	(1.012)	(0.452)
REPETIT	0.024	0.018	0.023	−0.041
	(0.978)	(0.041)	(0.864)	(0.089)
PHYEFF	−0.674	−0.971	0.006	−0.997**
	(1.556)	(1.571)	(0.208)	(2.240)
MONTHS	−0.0012	−0.0006	−0.00003	0.002
	(0.218)	(0.022)	(0.193)	(0.651)
Economic				
UNION	0.074	0.529	0.074	0.090
	(0.822)	(0.939)	(0.782)	(0.939)
UNFAIRPR	0.475	0.318	−0.0004	−0.054
	(0.309)	(0.466)	(0.003)	(0.466)
PAIDVAC	0.736	0.535	−0.172*	−0.091
	(0.891)	(0.860)	(1.753)	(0.860)

continued overleaf

Table 2—*continued*

	Estimated coefficient and absolute *t*-statistic			
	1	2	3	
			Two-part model	
			3a	3b
	OLS, dependent variable is ABSENT14	Two-limit Tobit, dependent variable is ABSENT14	Probit, dependent variable is BINARYABS	OLS, dependent variable is ABSENT14
MATERPY	0.924	1.282	0.243	0.184
	(0.768)	(1.129)	(1.495)	(1.129)
COMMUTE	0.032	0.0176	0.0024	0.003*
	(1.477)	(1.251)	(0.970)	(2.250)
MOONLITE	0.914	0.023	−0.081	−0.106
	(0.620)	(0.808)	(0.677)	(0.808)
PAIDSKLV	−0.820	−0.635	−0.165*	−0.108
	(0.915)	(1.024)	(1.831)	(1.024)
SHORTAGE	0.375	0.223	−0.044	−0.038
	(0.444)	(0.426)	(0.503)	(0.426)
WAGE	−0.0001	−0.00005	−0.0001	−0.000009
	(0.729)	(0.826)	(1.070)	(0.826)
OTHERINC	−0.00002	−0.000006	−0.000002	−0.000001
	(0.096)	(0.128)	(0.037)	(0.138)
Sample size	1308	1308	1308	222
R^2	0.102			0.415
−2 Ln (likelihood) sig:		0.0001	0.0001	

*Indicates significance at 0.05 level in one-tailed test.
**Indicates significance at 0.01 level in one-tailed test.

the one hand, employees in a large firm may feel they can be more easily replaced for a day or two since there are many additional co-workers to pick up the slack who would not be available in a small firm [36]. On the other hand, large firms may attract more committed workers who are less prone to absenteeism [37, 38].

RESULTS

Table 2 presents the results from Ordinary Least Squares regressions (OLS or MR), two-limit Tobit regressions, and a Probit regression. Four different regressions are reported: column 1 reports OLS (MR) results explaining ABSENT14 among the entire sample of 1302; column 2 reports two-limit Tobit results explaining ABSENT14 among 1302 subjects; column 3 reports Probit results explaining BINARYABS for 1302 subjects; and OLS results explaining ABSENT14 are presented in column 4 for 222 subjects with one or more absence days.

The numbers in Table 2 that do not have parentheses are estimated partial derivatives. For Ordinary Least Squares and Tobit, the coefficients are untransformed slope estimates. Partial derivatives must be derived from Probit parameter estimates [29]. The estimated partial derivatives indicate how a 1-unit change in an independent variable is associated with a change in the dependent variable, other things being equal. The numbers in parentheses are estimated asymptotic (Tobit, Probit) and finite sample (OLS) *t*-statistics.

Consider first the similarities and dissimilarities of the three estimating techniques. Columns 1 and 2 provide similar results. With only one exception, the same variables achieve statistical significance in both columns: being female with young children (FEMALE∗KIDS < 5), reporting sleep trouble

Table 3. Tests for the importance of groups of variables

Technique: Ordinary least squares (multiple regression)
Dependent variable: Absences during the past 14 days (ABSENT14)

Independent variables included (×) and excluded (○) from the regressions	1 *a*	1 *b*	2 *a*	2 *b*	3 *a*	3 *b*	4 *a*	4 *b*
Demographic AGE through FEMALE∗KIDS < 5	○	×	×	×	×	×	×	×
Health COLDFLU through OCCDTH	×	×	○	×	×	×	×	×
Job aspects OVERTM through MONTHS	×	×	×	×	○	×	×	×
Economic UNION through OTHERINC	×	×	×	×	×	×	○	×
	1		2		3		4	
F, comparing regression *a* with *b*	7.19*		12.86*		6.50*		4.88*	

*Statistically significant at the 0.01 level.

(SLEEP), being excessively overweight (FAT), and the three hazardous working conditions variables (DANGER, INJURY, OCCDTH). The one exception is for having flexibility in work hours (SAMEHRS), which achieves statistical significance in Tobit, but not in the Ordinary Least Squares or Multiple Regression results. Moreover, the estimated partial derivatives bear strong resemblance across the first two columns, especially for the variables that achieve statistical significance.

While the results in column 3a bear some resemblance to those in columns 1 and 2, the results in column 3b do not. Similarities between 3a on the one hand, and 1 and 2 on the other include the statistically significant results for being female with small children (FEMALE∗KIDS < 5), holding a dangerous job (DANGER, OCCDTH), and recently suffering an injury or illness caused or aggravated by the job (INJURY). The similar results for 3b compared to 1 and 2 pertain to being female with young children (FEMALE∗KIDS < 5), reporting sleep trouble (SLEEP), and holding a dangerous job (DANGER). But, there are striking dissimilarities between 3b on the one hand, and 1 and 2 on the other. Age, age squared, married with spouse present (MARRIED), reporting a cold or flu (COLDFLU), suffering back pain (BACKPAIN), holding a job requiring considerable physical effort (PHYEFF), and commuting distance (COMMUTE) are statistically significant in column 3b, but not in columns 1 and 2.

Table 3 presents results on the overall significance of categories of variables. Long and short OLS regressions are run which compare sum of squared residuals. Long regressions include all 37 covariates, while short regressions exclude one subset of covariates. The " × s" indicate which set of covariates are included, and "○s" indicate omitted variables. The largest F-statistic is generated by the health variables, and the smallest by the economic variables.

DISCUSSION

Comparing OLS, Tobit and two-part models

The conflicting results comparing 3a or 3b with 1 and 2 can be explained by the contrasting results between 3a and 3b. Prevalence of a cold or flu does not achieve statistical significance in either columns 1 or 2. The COLDFLU variable does achieve statistical significance in 3a and 3b, but the signs are reversed. A positive association is present in 3a, while a negative association is present in 3b. Apparently, the negative association among those with positive absences (3b) cancels with the positive association between COLDFLU and BINARYABS among all subjects in the sample. The results are easily explained. Colds and flu are frequently responsible for 1 or 2 days of absence, but seldom four or more. Four or more days of absence are most frequently the result of serious sickness or injury. Similar cancelling effects are apparent for smoking (SMOKE), back pain, working the same hours with no flexibility (SAMEHRS), jobs requiring physical effort (PHYEFF), and commuting (COMMUTE).

The statistically insignificant results on age and age squared in columns 1, 2 and 3a, and significant results in column 3b are revealing. Older workers are apparently *not* more likely than younger workers to report absence spells. But, once absent, older workers tend to stay away longer than younger workers, although the relationship increases at a decreasing rate, i.e. the sign on age squared is negative in column 3b. It is not surprising that older workers should stay out longer than younger workers, given that immune response decreases with age.

Rather than attempting to suggest which is the "best" model, more insight appears to be gained by treating the models as complementary, especially when comparing either a single equation model with the two-part model. The single equation models (columns 1 and 2) provide a broad macro-view and indicate covariates significantly associated with the total amount of absenteeism. The two-part model provides a micro-view and indicates first the covariates associated with any absences, and second the number of absences among those with at least one. If firms are concerned with minimizing long absences per employee, model 3b should be considered. This may be especially relevant if these absent employees occupy key jobs which are difficult to fill with temporary help. There may be increasing marginal costs associated with days absent. If firms are concerned with minimizing total absences, they should prefer models in columns 1 and 2. If there are rapidly decreasing marginal costs associated with absences, model 3a may be preferred.

The results comparing OLS (column 1) with Tobit (column 2) suggest ignoring censoring results in parameter estimates that are biased toward zero. Virtually all of the significant Tobit estimates are greater than the OLS estimates. The tendency for OLS to underestimate is repeatedly observed in the literature [30].

Demographic variables

Demographic variables such as age, race, gender, years of schooling, and marital status have received attention from researchers in psychology, occupational medicine, and economics [16, 17, 21, 39–43]. It is encouraging to note that for some of these variables, similar results were obtained by researchers from different fields. Thus, the economists Allen [16, 17] and Leigh [21] find education to be negatively associated with absenteeism; while the epidemiologists, Taylor [41] and Pines *et al.* [42], find social and occupational standing (which is strongly correlated with education) also to be negatively associated. But Taylor and Pines *et al.* were limited to samples of hospital employees and U.K. post office staff. Pines *et al.*, Allen, and Leigh also find marriage to lower absences for men; while marriage with the presence of small children raises absences for women. The results for race are mixed. Using data from one plant of a large manufacturing company in Southeast England, Baker and Pocock [43] find Caucasians to report the fewest number of absences, while Asians and West Indians report the greatest. Leigh, Allen, and Sexton and Schumann [44] do not find differences between whites and non-whites using U.S. data. The evidence for age is also mixed. Using data from Israel, Pines *et al.* find that advancing age increases absenteeism. Using data from Maryland state employees, Sexton and Schumann find age decreases the frequency of absences. Leigh and Allen do not find any consistent age effect.

The results from Table 2 suggest that only age and being a female with small children could be regarded as possible demographic predictors of absenteeism. None of the other demographic variables achieved statistical significance in columns 1 through 3b, or in additional regressions available from the author which combined the demographic variables with each of the other three categories. The results for the female with young children interaction variable are encouraging in that they are consistent with earlier findings in Pines *et al.* [42] and Leigh [21]. Both the t-statistics (2.9, 3.7, 1.98, 2.88), which indicate statistical significance and the sizes at the estimated partial derivatives (2.4, 3.4, 0.83, 1.126) are among the highest for any of the 37 independent variables. Mothers with small children are more likely to report absences than men or than women with no small children. These results have a clear policy implication. If firms and organizations want to decrease absences among their female employees with small children, they should consider sponsoring day care centers, or supporting legislation to provide day care, or subsidies for day care [45]. As day care becomes more available and fathers begin taking greater responsibility for raising children, the interaction variable may lose its importance.

The statistically insignificant results for education were unexpected. Education is one of the strongest predictors of health and longevity [46–48]. But few, if any, prior studies of education-health associations enter as many covariates as entered in this study. It could be that the strong education–health associations found in other studies are the result of education's correlations with income, occupation, smoking, drinking, disability, back pain, obesity, and so on. It could also be that absence from the job is more than a measure of health [49].

Married, spouse present (MARRIED) achieves statistical significance only in column 3b. The negative sign on the coefficient in 3b suggests that while marital status is not associated with number of absence spells, once an absence occurs, people who are married with spouse present tend to report back to work sooner than divorced, widowed, never married, and separated people. The quick return of married, spouse present people could be the result of home medical care provided by a loving spouse [50, 51].

Health variables

In studies by economists Leigh [21] and Paringer [20] and the epidemiologist Taylor [41], overall measures of respondents' health appeared to be the most important covariate of absenteeism. The evidence suggests that a large, if not the major, amount of absenteeism is the result of genuine illness or injury, and not malingering. Steers and Rhodes [52] state that questions involving the health and the "ability to attend, remain an area deserving further study; little has, in fact, been done on the topic". The importance of health is underscored by the results in Table 3. The results from Table 3 suggest that the health variables, as a group, explain more variation in absenteeism, adjusting for degrees of freedom, than any of the other four sets of covariates. These results are consistent with those in Paringer in which health variables were found to be more strongly associated with absenteeism than economic variables [20].

While the list of health variables obviously does not include all possible behaviors and afflictions, the list, nevertheless, appears to be one of the longest in the absence literature. The variable which has received the most attention in prior studies has probably been smoking. Parkes [14] reviewed these studies, most of which were based on regional samples. The QES national sample would be preferred over regional samples since there are great differences in smoking behavior across regions in the U.S. The prevalence of *current* cigarette smoking varies more than two-fold, from 14.7% in Utah to 34.4% in Kentucky [53]. The QES variable which acts as a proxy for smoking, SMOKE, suffers a number of drawbacks, however. SMOKE is a binary variable and does not distinguish among cigarettes, pipes or cigars. Moreover, SMOKE does not indicate how much is smoked or whether the person recently quit. That is, a person who recently quit would be included in the nonsmoker category. The poor quality of SMOKE may, thus, explain the disappointing results.

The overweight variable (FAT) is included to account for the effects of excessive weight on absenteeism reviewed in Parkes [15]. Parkes finds a "U" shape relation between relative weight and absenteeism. Skinny and fat people report more absences than persons of average weight for their height. This study considers only the fat side of the curve. Evidence in the QES suggests that overweight persons reported roughly 1 to 1.3 more absence days than other persons. But, the insignificance of FAT in column 3b suggests that while overweight people have more frequent absences, they tend to be short absences.

Few absence investigators consider whether the respondents reported a cold or flu during the past year, had insomia, amount of drinking, poor appetite, or disability as possible correlates of absenteeism. The evidence in Table 2 suggests that insomnia (SLEEP) may be an important predictor of absenteeism since estimated coefficients on SLEEP are satistically significant at better than the 0.05 level in a one-tailed test in columns 1, 2 and 3b. Persons who report no sleep trouble indicate roughly 1.4 fewer absent days than persons who do report sleep trouble. This appears to be a unique finding, apart from a brief remark in Jenkins [9]. But Jenkins' sample was limited to executive offices in the English Civil Service. A literature search did not uncover any prior studies on possible correlations between sleep trouble and absenteeism. No references to sleep or insomnia appear in Goodman *et al.* [54], Dilts *et al.* [55] or Chadwick-Jones *et al.* [56].

Andersson [57] finds evidence that back pain is the most frequently cited reason for absences. The evidence in the QES partially supports these findings. Persons who report back pain also report 0.948 more days absent in column 3b for any given absence spell. The result is statistically significant at the 0.05 level in a one-tailed test in column 3b. Unfortunately, BACKPAIN does not achieve significance in the other columns, thus inviting criticism of its importance in having an independent overall association with absenteeism.

Alcohol use was not found to be associated with absences. This is consistent with the finding of strong associations between alcohol use and earnings in other studies using these data [25, 58].

Poor appetite and reporting a work-limiting disability are not associated with either the number of absence spells or the length of absences. The insignificance of poor appetite is in contrast to results found in Jenkins [9]. But, Jenkins' data were based on a regional sample of nurses in London.

Variables reflecting the hazardous nature of the job—the respondents' evaluation of dangerous conditions (DANGER), the respondents' experiences with job-related injuries or illnesses (INJURY), and the probability of dying on the job (OCCDTH)—consistently draw among the largest t-statistics and highest levels of significance and largest estimated partial derivatives of all the possible covariates. Sizes of estimated partial derivatives and of t-statistics on DANGER, INJURY, and OCCDTH, in fact, rival those on the top two variables, the FEMALE∗KIDS interaction and SLEEP. These results are remarkable considering that DANGER, INJURY, and OCCDTH are undoubtedly highly correlated among themselves. When explanatory variables are correlated, multicollinearity results. Multicollinearity shrinks t-statistics and levels of statistical significance because the Tobit and OLS procedures face greater obstacles in estimating the independent association of any one explanatory variable with the dependent variable when multicollinearity is high than when it is low [35].

Persons who work at dangerous jobs (DANGER) report between 1.4 and 1.6 more absence days than persons in safe jobs. Persons who have an injury or illness they believe was caused or exaserbated by the job report between 1.1 and 1.5 more absence days than others. Persons who report one-half unit increase (from 1/1000 to 1.5/1000) in the probability of death on the job also report a total of about 1.6–2.4 more absence days. A one-half point increase is substantial given the OCCDTH mean of 0.052.

Job aspects

Perhaps the single question that has received more attention by psychologists than any other concerns the effects of job satisfaction on absenteeism [59]. Muchinsky's [1] widely cited review concludes that 'highly consistent' results between job satisfaction and lower absenteeism had been found in prior research. Curiously, since 1977, the consensus has dissolved as evidenced in a comment by Keller [60] that "more recent research has been almost as consistent in finding that job satisfaction was unrelated to absenteeism . . .". But, most of the psychological studies have been limited to simple univariate correlations between satisfaction and absenteeism.

Again, because of the unique data available in the QES, the association of job satisfaction independent of wages, schooling, age, race, gender, and so on, could be assessed.

The results are disappointing for the single job satisfaction hypothesis. The binary job satisfaction variable is not statistically significant in any of the models. These results should not be interpreted as suggesting that job satisfaction in the global sense is unimportant. A more sympathetic reading of the results indicates that it is important to go beyond the simple univariate associations and ask what contributes to job satisfaction. It is likely, for example, that job satisfaction is related to excessive overtime, inflexible hours, hazardous nature of the job, size of the organization, months of tenure with the job, repetitive nature of the job, and so on.

The inflexible hours variable, SAMEHRS, appears to be a predictor of the number of absence spells. It is significant in columns 2 and 3a. Moreover, the estimated partial derivatives indicate that persons with inflexible hours report roughly 1.3 more absent days than persons with flexible hours. Moving to flexible hours may be a direct method whereby firms can achieve a lower absence rate, at relatively little cost. Moreover, this may be an especially effective device for lowering absences for workers with small children [45].

Economic variables

Economists Winkler [18] and Allen [16, 17] have suggested that wages and the availability of sick leave may play a role in explaining absenteeism. If sick leave is available, workers may be more inclined to report absences than if it is not. If sick leave is not available and workers are paid a high wage, then they may be less inclined to be absent than if they are paid low wages. The wage represents the 'opportunity cost' for taking an absence if sick leave is unavailable. The higher the wage, the greater the 'opportunity cost' of taking an absence. The results in Table 2 do not support the simple sick leave or wage hypothesis.

Six additional economic variables are considered. Seamonds [61] suggests that the threat of unemployment ought to decrease absenteeism. Leigh [19] finds that measures of unemployment are inversely associated with absenteeism, but the unemployment data used are crude since they applied to county- and industry-wide divisions. What should matter in the economic view is the individual's prospects for unemployment. The effect of the perception of unemployment on absenteeism is directly assessed with the QES which contains information on workers' subjective evaluations of unemployment in their area for their skills—SHORTAGE. The results do not support the unemployment hypothesis in any column.

Moonlighting, although never investigated by economists, may also play a role in explaining absences. It may be that the second job occasionally requires the worker to take time off from his first job. But again, the results are inconclusive. MOONLITE never achieves statistical significance. It could be that persons who choose a moonlighting job are a hearty group so that self-selection bias may explain the results.

Taylor and Pocock [62] find that excessive commuting distance is positively associated with absences. They hypothesize that people with long commutes would reap greater benefits from being absent since their total time away from home would be so much more than persons with short commutes. But, Taylor and Pocock limit their sample to only London office workers. Costs of commuting may dramatically differ across many cities as are available in the QES. It could be, on the other hand, that persons who commute a long distance are paid more and are more committed to their jobs. The results in Table 2 on COMMUTE support the Taylor and Pocock conclusion of a positive association between commuting and absenteeism for number of absence spells, but not length of absences.

The paid maternity leave variable, MATERPY, is included to further investigate the importance of small children in affecting female absences. The results are disappointing. No association is discovered. In part, these (lack of) statistical associations may be the result of only 18 people in the MATERPY category ($18 = 1304 \times 0.013$). Again, these data are from 1973. The number of workers who qualify for paid maternity leave today is much greater [45].

Whether the availability of paid vacation is important or not is considered for economic reasons. A worker who has paid vacation may feel less of a need to take absences than workers without a paid vacation. Alternatively, it could be that firms which pay for vacations attract exceptionally motivated workers, just as Garen [37] argued that large firms offering high wages attract hard-working employees. Both economic hypotheses are rejected. The results in Table 2 reveal insignificant estimated coefficients on the paid vacation variable.

If absences can be viewed as 'multi-vacations' from work, the economic theory outlined in Allen [16, 17] suggests that as the workers' non-job income rises, he or she will 'purchase' more vacation time. The evidence from Table 2 does not provide support for the hypothesis since the OTHERINC estimated partial derivative does not achieve statistical significance in any column.

CONCLUSION

Six conclusions can be drawn from this study. (1) The variable for mothers with small children ranks as the most important predictor among workers' personal characteristics. (2) Among the health variables, insomnia, being overweight, and hazardous working conditions appear to be the most important. (3) Inflexible hours is apparently the most important job-related variable. (4) A number of frequently discussed predictors including race, unionization, wages and job satisfaction flunked their statistical significance tests. (5) Health variables and hazardous conditions at the job appear to be among the most important, yet least researched predictors of absenteeism. (6) Single equations, Tobit regressions, and two-part models are complementary research techniques in analyses of absenteeism.

Acknowledgements—The author acknowledges help from Harold Waldon, Roger Folsom and Pamela Davis. The research was funded by grant no. R01 OH02586 from the National Institute for Occupational Safety and Health.

REFERENCES

1. Muchinsky P. N. Employee absenteeism: a review of the literature. *J. Vocational Behav.* **10**, 316–304, 1977.
2. Steers R. M. and Rhodes S. R. Major influences on employee attendance: a process model. *J. appl. Psychol.* **63**, 391–407, 1978.
3. Breaugh J. A. Predicting absenteeism from prior absenteeism and work attitudes. *J. appl. Psychol.* **66**, 555–560, 1981.
4. Hammer T. H. and Landau J. Methodological issues in the use of absence data. *J. appl. Psychol.* **66**, 574–581, 1981.
5. Baba V. V. Methodological issues in modeling absence: a comparison of least squares and Tobit analyses. *J. appl. Psychol.* **75**, 428–432, 1990.
6. Brooke P. P. and Price J. L. The determinants of employee absenteeism: an empirical test of a causal model. *J. occupat. Psychol.* **62**, Part 1, 1–20, 1989.
7. Clegg C. W. Psychology of employee lateness, absence and turnover: a methodological critique and an empirical study. *J. appl. Psychol.* **68**, 88–101, 1983.
8. Hazcoun R. R. and Dupont S. An empirical analysis of the activity patterns of employees who are absent. *Relations Indust.* **43**, No. 1, 153–166, 1988.
9. Jenkins R. Minor psychiatric morbidity in employed young men and women and its contribution to sickness absence. *Br. J. Indust. Med.* **42**, 144–154, 1985.
10. Farrell D. and Stamm C. L. Meta-analysis of the correlates of employee absence. *Hum. Relations* **41**, 211–227, 1988.
11. Athanasou J. A. Sickness absence and smoking behavior and its consequences: a review. *J. occupat. Med.* **17**, 441–445, 1975.
12. Bass F. Epidemiology of cigarettes: usage, illness, and cessation. *Behav. Med. Update* **2**, 13–16, 1980.
13. Janzon L., Lindell S. F. and Trell R. S. Smoking and disease attitudes and knowledge in middle-aged men. *Scand. J. Soc. Med.* **9**, 127–133, 1981.
14. Parkes K. P. Smoking as a moderator of the relationship between affective state and absence from work. *J. appl. Psychol.* **68**, 698–708, 1983.
15. Parkes K. P. Relative weight, smoking, and mental health as predictors of sickness and absence from work. *J. appl. Psychol.* **72**, 275–286, 1987.
16. Allen S. G. An empirical model of work attendance. *Rev. Econom. Stat.* **63**, No. 1, 77–87, 1981a.
17. Allen S. G. Compensation, safety and absenteeism: evidence from the paper industry. *Industrial Labor Relations Rev.* **34**, 207–218, 1981b.
18. Winkler D. The effects of sickleave policy on teacher absenteeism. *Indust. Labor Relations Rev.* **33**, 232–240, 1980.
19. Leigh J. P. The effects of unemployment and the business cycle on absenteeism. *J. Econom. Business* **37**, 170–195, 1985b.
20. Paringer L. Women and absenteeism: health or economics? *Am. Economic Rev.* **73**, 123–127, 1983.
21. Leigh J. P. Correlates of absence from work due to illness. *Hum. Relations* **34**, 81–100, 1986.
22. Leigh J. P. Assessing the importance of an independent variable in multiple regression: is stepwise unwise? *J. clin. Epidemiol.* **41**, No. 7, 669–677, 1988.
23. Goldberger A. S. and Jochens D. B. Note on stepwise least squares. *J. Am. Stat. Assoc.* **49**, No. 2, 105–110, 1961.
24. Kallenberg A. L. *America at Work: National Survey of Employees and Employers.* Social Science Research Council, New York, 1986.
25. Cook P. J. The social costs of drinking. Expert Meeting on Negative Social Consequences of Alcohol Use, Oslo, Norway, August, 1990.

26. Fennell M. L., Rodin M. B. and Kantor G. K. Problems of the work setting, drinking and reasons for drinking. *Soc. Forces* **60**, No. 1, 114–132, 1981.
27. Robinson J. C. Workplace hazards and workers' desires for union representation. *J. Labor Res.* **9**, 237–250, 1988.
28. Leigh J. P. New estimates of mortality rates in 347 occupations. *J. occupat. Med.* **29**, No. 6, 510–519, 1987.
29. Maddala G. S. *Limited Dependent and Qualitative Variables in Econometrics*. Cambridge University Press, Cambridge, 1983.
30. Greene W. H. *Econometric Analysis*, Chap. 21, pp. 730, 738. Macmillan, New York, 1990.
31. Kalbfleisch J. D. and Prentice R. L. *The Statistical Analysis of Failure Time Data*. Wiley, New York, 1980.
32. Vijverberg W. P. M. Consistent estimates of the wage equation when individuals choose among income-earning activities. *Southern Econom. J.* **52**, No. 4, 1028–1042, 1986.
33. Greene W. H. *LIMDEP, An Econometric Software Package*. New York University, New York, 1989.
34. Hay J. W., Len R. and Rohrer P. Ordinary least squares and sample-selection models of health care demand. *J. Business Economic Stat.* **5**, No. 4, 499–506, 1987.
35. Maddala G. S. *Introduction to Econometrics*. Macmillan, New York, 1988.
36. Porter L. W. and Lawler E. E. Properties of orgnanization structure in relation to job attitudes and job behavior. *Psychol. Bull.* **64**, No. 1, 23–51, 1965.
37. Garen J. Worker heterogeneity, job screening, and firm size. *J. Polit. Economy* **93**, 715–739, 1985.
38. Brown C., Hamilton J. and Medoff J. *Employers Large and Small*. Harvard University Press, Cambridge, MA, 1990.
39. Nicholson N., Brown C. A. and Chadwick-Jones J. K. Absence from work and personal characteristics. *J. appl. Psychol.* **62**, 319–327, 1977.
40. Ortmeyer C. G. Variations on mortality, morbidity, and health care by marital status. In *Mortality and Morbidity in the United States* (Edited by Erhardt C. L. and Berlin J. E.), pp. 159–588. Garvard University Press, Cambridge, MA, 1974.
41. Taylor P. J. Occupational and regional association of death, disablement, and sickness absence among post office staff, 1972–75. *Br. J. Indust. Med.* **33**, 203–235, 1976.
42. Pines A., Skulkao K., Pollak E., Peritz E. and Steif J. Rates of sickness absenteeism among employees of a modern hospital. *Br. J. Indust. Med.* **42**, 326–335, 1985.
43. Baker L. and Pocock S. J. Ethic differences in certified sickness absence. *Br. J. Indust. Med.* **39**, No. 2, 277–282, 1982.
44. Sexton M. and Schumann B. C. Sex, race, age and hypertension as determinants of employee absenteeism. *Am. J. Epidemiol.* **122**, No. 2, 302–310, 1985.
45. Fuchs V. R. *Women's Quest for Economic Equality*. Harvard University Press, Cambridge, MA, 1988.
46. Grossman M. The correlation between health and schooling. In *Household Production and Consumption—Studies in Income and Wealth* (Edited by Terleckyj N. E.). Columbia University Press, New York, 1975.
47. Kenkel D. S. Health behavior, health knowledge, and schooling. *J. Political Economy*. **99**, 287–305, 1991.
48. Pincus T. Formal education level: a marker for the importance of behavioral variables in the pathogenesis, morbidity, and mortality of most diseases? *J. Rheumatol.* **15**, No. 10, 1457–1460, 1988.
49. Butler J. S., Burkhauser R. V., Mitchell J. M. and Pincus T. P. Measurement error in self-reported health variables. *Rev. Econom. Stat.* **69**, No. 4, 644–650, 1987.
50. Taubman P. and Rosen S. Healthiness, education, and marital status. In *Economic Aspects of Health* (Edited by Fuchs V. R.). University of Chicago Press, Chicago, 1982.
51. Berger M. C. Labor supply and spouse's health. *Soc. Sci. Q.* **64**, No. 3, 494–509, 1983.
52. Steers R. M. and Rhodes R. S. Knowledge and speculation about absenteeism. In *Absenteeism* (Edited by Goodman P. S. *et al.*). Jossey-Bass, San Francisco, 1984.
53. Editor. Behavioral risk factor surveillance. *MMWR* **39**, No. S52, 1–22, 1988.
54. Goodman P. *et al.* (Eds) *Absenteeism*. Jossey-Bass, San Francisco, 1984.
55. Dilts D. A., Deitsch C. R. and Paul R. J. *Getting Absent Workers Back on the Job*. Quoram Books, Westport, CT, 1985.
56. Chadwick-Jones J. K., Nicholson N. and Brown C. *Social Psychology of Absenteeism*. Praeger, New York, 1982.
57. Anderson G. B. J., Svensson H. and Oden A. The intensity of work recovery in low back pain. *Spine* **8**, No. 6, 880–884, 1983.
58. Berger M. C. and Leigh J. P. The effect of alcohol use on wages. *Appl. Econ.* **20**, 1343–1351, 1988.
59. Nicholson N. Industrial absence as an indicator of employee motivation and job satisfaction: a study of contrasting technologies in the North of England and South Wales. Unpublished doctoral dissertation, University of Wales, 1975.
60. Keller R. T. Predicting absenteesism from prior absenteeism, attitudinal factors and nonattitudinal factors. *J. appl. Psychol.* **68**, 536–540, 1983.
61. Seamonds B. C. Stress factors and their effect on absenteeism in a corporate employee group. *J. occupat. Med.* **24**, 393–397, 1982.
62. Taylor P. T. and Pocock S. J. Commuter travel and sickness absence of London office workers. *Br. J. Prevent. soc. Med.* **26**, 165–172, 1972.