ILR Review

Volume 61 | Number 1 Article 6

October 2007

New Evidence on the Determinants of Absenteeism Using Linked Employer-Employee Data

Georges Dionne
HEC Montréal

Benoit Dostie
HEC Montréal

New Evidence on the Determinants of Absenteeism Using Linked Employer-Employee Data

Abstract
This paper provides new evidence on the determinants of absenteeism. The authors extend the typical labor-leisure model used to analyze the decision to skip work to include firm-level policy variables relevant to the absenteeism decision and uncertainty about the cost of absenteeism. Estimates based on data from Statistics Canada's Workplace Employee Survey (1999–2002), with controls for observed and unobserved demographic, job, and firm characteristics (including workplace practices), indicate that work arrangements were important determinants of absence. For example, the authors find strong evidence that standard weekday work hours, work-at-home options, and reduced workweeks were associated with reduced absence, whereas shift work and compressed work weeks were associated with increased absence.

Keywords
absenteeism

NEW EVIDENCE ON THE DETERMINANTS OF ABSENTEEISM USING LINKED EMPLOYER-EMPLOYEE DATA

GEORGES DIONNE and BENOIT DOSTIE*

This paper provides new evidence on the determinants of absenteeism. The authors extend the typical labor-leisure model used to analyze the decision to skip work to include firm-level policy variables relevant to the absenteeism decision and uncertainty about the cost of absenteeism. Estimates based on data from Statistics Canada's Workplace Employee Survey (1999–2002), with controls for observed and unobserved demographic, job, and firm characteristics (including workplace practices), indicate that work arrangements were important determinants of absence. For example, the authors find strong evidence that standard weekday work hours, work-at-home options, and reduced workweeks were associated with reduced absence, whereas shift work and compressed work weeks were associated with increased absence.

In this paper, we provide new evidence on the determinants of absenteeism using linked employer-employee data. It has long been recognized that an individual's decision to skip work might be influenced by features of the firm's personnel policies or organizational structure (Frankel 1921). Linked data thus provide a unique opportunity to sort out the different causes of absenteeism. Despite absenteeism's rising frequency and associated cost (Akyeampong 2005), relatively few studies have investigated its determinants. Moreover, it could be argued that most existing studies of the determinants of absenteeism suffer from the use of less than adequate data. A first strand of the literature focuses on only one kind of absenteeism, namely absenteeism due (officially) to health reasons. These studies generally use data from health insurance companies or government agencies.[1] A second strand of the literature uses detailed absenteeism data from one company or a very small sample of firms.[2] It is not

*Georges Dionne is Professor of Finance and Canada Research Chair in Risk Management, and Benoit Dostie is Associate Professor of Economics, both at HEC Montréal. The authors are grateful to seminar participants at the "Journées du CIRPÉE" (Knowlton, Québec), SCSE (Manoir Richelieu, Québec), CEA (Hamilton, Ontario), and SOLE (San Francisco), and in particular to André-Pierre Blanchard-Dionne, Cindy Carter, C. Robert Clark, Pierre Fortin, Pierre Lefebvre, Mathieu Maurice, and Kathryn Shaw. They also thank Lene Kromann for excellent research assistance.

A data appendix with additional results, and copies of the computer programs used to generate the results presented in the paper, are available from the second author at Institute of Applied Economics, HEC Montréal, 3000, chemin de la Côte-Sainte-Catherine, Montréal, H3T 2A7; benoit.dostie@hec.ca.

[1]For example, Henrekson and Persson (2004) used aggregate data from the National Social Insurance Board of Sweden, and Johansson and Palme (2002) used data from the 1991 Swedish Level of Living Survey (SLLS). In the United States, Vistnes (1997) used the 1987 National Medical Expenditure Survey.

[2]Kauermann and Ortlieb (2004) used absenteeism data from one German firm. Similarly, Barmby (2002) used data on only one U.K. manufacturing firm. Delgado and Kniesner (1997) focused on London bus operators. Drago and Wooden (1992) worked on a sample of 15 firms in the United States, Canada, and New-Zealand. Barmby, Orme, and Treble (1991) used data on four factories of an unidentified firm. Wilson and Peel (1991) analyzed data on a sample of 52 firms in the engineering and metal industry in the United Kingdom. Dunn and Youngblood (1986) used 1977 data from one utility company.

Industrial and Labor Relations Review, Vol. 61, No. 1 (October 2007).
0019-7939/00/6101 $01.00

clear that the results from these studies are generalizable outside their small samples.[3]

Our work is more closely related to the second strand of the literature than to the first. However, we examine the determinants of absenteeism using survey data, the Workplace Employee Survey (WES) 1999-2002 from Statistics Canada. The WES has numerous advantages for studying the determinants of absenteeism: (1) the survey is designed to be representative of the whole universe of workplaces operating in Canada; (2) in each sampled workplace, a subset of workers from the firm is sampled, so that the survey is also representative of the universe of workers in Canada;[4] (3) since the survey generates linked employer-employee information, we have detailed micro data on each of these workers, including days of absence during the year, demographic and job characteristics, preferences, and human capital variables (this is in addition to the usual firm-level characteristics); (4) each worker is asked to recall the number of days absent from work in the past year; (5) the linked nature of the data allows us to take into account unobserved firm heterogeneity; and (6) the longitudinal nature of the data allows us to take into account unobserved worker heterogeneity.

We start by extending the typical labor-leisure model for analyzing the decision to skip work so that it includes firm-level policy variables relevant to the absenteeism decision and uncertainty about the cost of absenteeism to the worker. We next describe an econometric model that explicitly accommodates the count nature of absenteeism data and also incorporates unobserved heterogeneity at both the individual and firm level. Following a brief description of the data sources and variables, we then proceed to an examination of the results and a brief conclusion.

Theoretical Framework

We start by setting up the typical labor-leisure choice model to study the absenteeism decision (see Allen 1981; Allen 1983; Barmby, Orme, and Treble 1991; Delgado and Kniesner 1997; and Dunn and Youngblood 1986).[5] We assume that each job offers a work schedule as well as a wage rate. Since search is costly, a worker may accept a job offer even though, at the contracted number of work hours (t^c), his marginal rate of substitution between leisure and income does not equal the wage rate (w). When a worker contracts for more than his desired hours given w, he retains an incentive to consume more leisure. One way of doing so is to be absent from work. In this theoretical framework, an emphasis will be placed on the explicit random cost of such a decision and on how workplace and job characteristics affect this decision. These two aspects, neither of which has been addressed in the literature, will become important in the empirical part of the paper.

Absenteeism results in lost output when the absent worker either is replaced by someone who is generally less efficient or is not replaced at all. For the employment relation to continue, the firm must be compensated for this loss. In addition to losing earnings he would have received if he had reported, the worker faces a penalty (D) for each scheduled work period missed. In practice, this penalty will be observed in the form of a decreased probability of receiving a promotion or merit wage increase and an increased likelihood of being dismissed. Denoting the desired time absent from work as t^a, one can then write

[3]A notable exception is Allen (1981), who used the 1972-73 Quality of Employment Survey. However, the Quality of Employment Survey does not have any information about the employer and therefore cannot be used to study the link between workplace practices and absenteeism.

Other papers examining absenteeism include Gilleskie (1998), which focused on the absenteeism decision of individuals with acute illnesses, Ehrenberg (1970), which studied the link between absenteeism and the decision of the firm to use overtime, and Allen (1983), which estimated the cost of absenteeism.

A third strand takes a more macroeconomic approach. For example, Kenyon and Dawkins (1989) used aggregate Australian time-series data.

[4]Abowd and Kramarz (1999) classified WES as a survey in which both the sample of workplaces and the sample of workers are cross-sectionally representative of the target population.

[5]The following discussion also draws from Vistnes (1997) and Johansson and Palme (1996).

$$D = D(t^a) \quad D' \geq 0, \; D'' \geq 0, \; D(0) = 0.$$

The workers who miss the most days pay the largest penalties. The costs to the firm of increased amounts of absenteeism are presumed to be non-decreasing, yielding a constant or graduated penalty structure. Workers with perfect attendance records are not penalized at all. Since the worker does not really know this potential cost when he makes his decision, we consider the possibility that $D(t^a)$ can be a random variable. We write $\tilde{D}(t^a)$ when this is the case.

Holding work schedule flexibility constant, the work attendance decision can be analyzed within the traditional labor-leisure choice framework. Workers maximize an expected utility function containing consumption (C) and total leisure time (L) as its arguments:[6]

$$(1) \qquad EU = EU(C, L; P, F).$$

The expected utility of the worker is also a function of a vector of personal characteristics (P) and a vector of firm characteristics (F). Letting R equal the individual non-labor income, the budget constraint of the worker is

$$(2) \qquad R + w(t^c - (1 - s_L)\, t^a) - \tilde{D}(t^a) = \tilde{C},$$

where the price of the consumption good C is normalized to one, t^c is contracted hours, w is the wage rate, and s_L is a variable that takes the value of one if a worker has full leave benefits and less than one otherwise.[7] Workers also face a time constraint of

$$(3) \qquad t - t^c - t^a - t^l = 0,$$

where t represents the total amount of time in the period under consideration and t^l is pure leisure time. We can thus write $t^a + t^l = L$. Substitution of (2) and (3) in (1) and differentiation of the latter with respect to t^a produces the first-order condition

$$(4) \qquad E[\,U_L - (w(1 - s_L) + D'(t^a))\,U_C] = 0,$$

where $U_k > 0$ indicates the partial derivative of U with respect to $k = L, C$. The variable $\tilde{D}(t^a)$ can be expressed more directly by defining w^a as the unit cost of being absent. Thus we can write $\tilde{D}(t^a) = \tilde{w}^a t^a$ and, as already mentioned, w^a can be a random variable when the decision on t^a is made. In this case, the first order condition (4) becomes

$$(5) \qquad E[\,U_L - (w(1 - s_L) + w^a)\,U_C] = 0.$$

A worker will be absent on any given day as long as the extra leisure is more valuable to him than the sum of the wages he would have earned that day and the resulting loss in future earnings. This means that the shadow price of time for absent workers is greater than the contracted wage.

By differentiating the first-order conditions for $s_L = 0$ and applying Cramer's Rule, one can show, under the usual conditions of a downward-sloping absenteeism demand curve, that

$$(6) \qquad \frac{\partial t^a}{\partial w} < 0,\; \frac{\partial t^a}{\partial R} > 0,$$
$$\frac{\partial t^a}{\partial t^c} > 0,\; \frac{\partial t^a}{\partial (Risk\ w^a)} < 0,$$
$$\frac{\partial t^a}{E(w^a)} < 0,$$

where $Risk\ w^a$ is a Rothschild and Stiglitz (1971) measure of the risk associated with w^a and $E(w^a)$ its mean. Details for the derivation of results in (6) are in an appendix available from the authors.

The effect of a change in the wage rate on time absent from work is ambiguous *a priori* because income and substitution effects operate in opposite directions. However, under the conditions of a downward-sloping absenteeism demand curve, a negative sign is obtained when s_L equals 0 or is sufficiently small. An increase in non-labor income leads to more demand for all non-inferior goods and services, including time absent from work. If the number of contracted hours changes, the number of absences moves in

---

[6] See Dionne and Eeckhoudt (1987) for an analysis of labor supply under uncertainty.

[7] As in Vistnes (1997), detailed information on sick leave provisions (stock of sick leave, carry-over provisions, whether sick leave benefits pay the worker fully or partially, and whether the sick leave can be applied toward early retirement or used for maternity leave) is not available. We do, however, have detailed information about work arrangements that might serve as proxies for these provisions.

the same direction because t^c generates a positive income effect. An increased average penalty for absenteeism reduces the number of days missed, as does an increased risk of penalty.

In cases where full leave benefits are available ($s_L = 1$), the product of w and t^a disappears from (2) and the first–order equilibrium condition becomes

$$E[U_L - w^a U_C] = 0. \tag{7}$$

In this particular case, the cost of absence is reduced to the resulting loss in future earnings. So the optimal decision of the worker involves a trade-off between the marginal benefit of leisure and future earnings or benefits. Unless the penalty function is made steeper, an individual will be absent more frequently in plants where sick leave is fully paid to absent workers. It should be noted that the effect of a wage change on the likelihood of absence is unambiguously positive in this case because there is no longer a substitution effect.

The model can be summarized as

$$t^a = t^a \; \underset{(-)}{(w,} \quad \underset{(+)}{R,} \quad \underset{(+)}{t^c,} \quad \underset{(-)}{E(w^a),} \quad \underset{(-)}{Risk\ w^a)} \tag{8}$$

We provide a structural form for these relationships in the next section.

Empirical Specification

From the above behavioral model, we can derive a structural econometric model of the absenteeism decision. Extending the model Hausman (1980) and Blomquist (1983) proposed for labor force participation, we can write the following functional form for the direct expected utility function of a given worker:

$$EU(C, t^a; P, F) = \exp\left\{-\left(1+\frac{\beta(C-\bar{P}-\bar{F})}{(-t^a+b))}\right)\right\}\left(\frac{-b+t^a}{\beta}\right), \tag{9}$$

where

$$\bar{P}+\bar{F} = P/\beta + F/\beta - \alpha/\beta^2$$

$$b = \alpha/\beta$$

$$C = R + wt^c - w(1-s_L)\,t^a -$$
$$E(w^a)\,t^a - (\beta/2(b-t^a))\,\sigma^2_{w^a}(t^a)^2.$$

α and β are parameters and β can be interpreted as the absolute risk aversion parameter. As shown by Hausman (1980), this specific utility function yields a linear relationship for the supply of hours of work and does not introduce restrictions on wage and income elasticities. Here we extend the model by considering a random variable that is assumed to follow a normal distribution with parameters $E(w^a)$ and $\sigma^2_{w^a}$. Hence we maximize the value of (9) with respect to t^a to obtain the desired linear relationship in w, R, $E(w^a)$, t^c, P, and F. As we will see below, a linear relationship is important for the econometric specification. From the expected utility function, one can verify that the demand for periods of absenteeism is equal to

$$t^a = \kappa(-\alpha(w(1-s_L) + E(w^a)) + \beta(R + t^c w) + P + F), \tag{10}$$

where $\kappa = 1/(1+\beta^2\sigma^2_{w^a})$. Under the assumption of normality for w^a, $\sigma^2_{w^a}$ is the measure of risk ($Risk\ w^a$). A positive α parameter and a positive β parameter are expected. Again, the results in (6) are verified with the particular expected utility function. In a more compact form, (10) can be rewritten as

$$t^a = \kappa(-\alpha w^* + \beta R^* + P + F), \tag{11}$$

where w^* can be interpreted as the relative cost of being absent and R^* as the virtual benefit or income related to absence.

In this simple model, where there is no unobserved heterogeneity, days of absence can be represented by a Poisson process. In fact, since absences are recorded as non-negative integers, modeling such data with a continuous distribution could lead to inconsistent parameter estimates. Let t^a_{ijt} be the observed number of days of absence for employee i in firm j at time t. The basic model is Poisson with parameter

$$\lambda_{ijt} = \exp\{\kappa(-\alpha w^*_{ijt} + \beta R^*_{ijt} + \gamma Z_{1it} + \eta Z_{2jt})\} > 0, \tag{12}$$

Table 1. Summary Statistics on Absenteeism in Canada, 1999.

Weeks Absent	*Frequency*	*Percentage*
0	9,717,342	90.16
1	669,090	6.21
2	185,702	1.72
3	25,927	0.24
4	31,343	0.29
5	7,437	0.07
6	22,676	0.21
7	14,159	0.13
8	15,538	0.14
9	10,675	0.10
10	3,986	0.04
Total	10,777,543	100.00

Note: Mean number of *days* absent (versus weeks absent): 3.691 (standard deviation, 6.665).

where Z_{1it} and Z_{2jt} are observed vectors of characteristics (of the worker and workplace, respectively). It should be repeated that t^c_{ijt} (desired contracted hours) is exogenous in the model. This decision variable is already fixed when the worker (or nature) makes a decision about t^a.

It is typical to introduce unobserved heterogeneity in the Poisson model through λ_{ijt} when we apply the model to a population of heterogeneous individuals and workplaces. We use the following parameterization for λ_{ijt}:

$$\lambda_{ijt} = \exp(\kappa(-\alpha w^*_{ijt} + \beta R^*_{ijt} + \gamma Z_{1it} + \eta Z_{2jt}) + \psi_j + \theta_{ij}). \tag{13}$$

The additional parameters ψ_j and θ_{ij} capture the impact of unobserved characteristics of the workplace and the worker, respectively.[8] These unobserved characteristics are assumed to be orthogonal to other observed characteristics. We assume both workplace and worker unobserved heterogeneity to be normally distributed with mean zero. The variance of $\psi_j(\sigma^2_\psi)$ is identified by the observation of many workers coming from the same workplace, while identification of the variance of $\theta_{ij}(\sigma^2_\theta)$ is possible by multiple observations of the same worker over time.[9]

Unobserved workplace heterogeneity might proxy for the cost of absence to the workplace when observed heterogeneity is not sufficiently informative. For example, the cost of absence to the firm might be fairly low if substitute workers are easily available and are as productive as regular workers (Allen 1983). Therefore, the econometrician might observe higher absenteeism than in an otherwise identical firm where such substitute workers are not available. From a statistical point of view, it is necessary to take into account both sources of heterogeneity in order to avoid the problem of spurious regressions due to multiple observations on the same worker over time and the same firm characteristics over its employees. Unobserved heterogeneity at the worker level might represent different preferences or work-ethic/motivation levels, or unobserved job characteristics like the safety of the work environment.

We use maximum likelihood methods to obtain estimates for the parameters, integrating out the two separate unobserved heterogeneity components. Since a closed form solution to the integral does not exist, the likelihood was computed by approximating the normal integral using a numerical integration algorithm based on Gauss-Hermite Quadrature. This algorithm selects a number of points and weights such that the weighted points approximate the normal distribution. Details for the derivation of the likelihood function are in an appendix available from the authors.

Data

We use data from the Workplace and Employee Survey (WES) 1999–2002 conducted by Statistics Canada. The survey is both longitudinal and linked in that it tracks the characteristics of the workers and of the workplaces over time. The target population

[8]Since we do not observe workers over different jobs, we cannot distinguish between unobserved heterogeneity at the level of the individual worker and at the level of the job.

[9]Note that this specification is not subject to the usual objection to the Poisson model, since the inclusion of unobserved firm and worker heterogeneity allows for over-dispersion at both the worker and firm level.

Table 2. Summary Statistics—Employees, 1999.

Variable	*Mean*	*Std. Dev.*
Demographic Characteristics		
Women	0.506	0.500
Black	0.012	0.109
Other Race	0.276	0.447
Married	0.568	0.495
Number of Pre-School-Aged Kids	0.228	0.554
Health		
No Activity Limitation	0.958	0.200
Human Capital		
High School Degree	0.178	0.382
Less Than Bachelor Degree	0.574	0.450
Bachelor Degree	0.130	0.337
Some Higher Education	0.054	0.092
Seniority	9.106	8.403
Experience	16.470	10.660
Income		
Income from Other Sources	2,120.118	11,226.875
Wage Contract		
Natural Logarithm of Hourly Wage	2.819	0.504
Contracted Hours	37.077	9.120
Work Arrangement		
Works Regular Hours	0.131	0.337
Usual Workweek Includes Saturday and Sunday	0.173	0.379
Work Flexible Hours	0.373	0.483
Does Not Work Traditional Hours (M–F between 6 a.m. and 6 p.m.)	0.667	0.471
Some Work Done at Home	0.275	0.446
Work Some Rotating Shift	0.058	0.233
Work on a Reduced Workweek	0.046	0.209
Work on Compressed Work Week Schedule	0.032	0.176
Covered by a Collective Bargaining Agreement	0.321	0.469
Technology		
Use Computer	0.627	0.484
Use Computer Assisted Design	0.130	0.337
Use Other Technology	0.249	0.432
Number of Employees	18,671	

for the "workplace" component of the survey is defined as all Canadian establishments that paid employees in March of the year of the survey. The survey, however, does not cover the Yukon, the Northwest Territories, or Nunavut. Establishments operating in fisheries, agriculture, and cattle farming are also excluded. For the "employee" component, the target population is all employees working, or on paid leave, in the workplace target population.

The workplace sample comes from the "Business Registry" of Statistics Canada, which contains information on every business operating in Canada. Employees are then sampled from an employee list provided by the selected workplaces. For every workplace, a maximum of twenty-four employees are selected, and for establishments with fewer than four employees, all employees are sampled. In the case of total non-response, respondents are withdrawn entirely from the survey and sampling weights are recalculated in order to preserve the representativeness of the sample. WES selects new employees and workplaces in odd years (at every third

Table 3. Summary Statistics—Workplace, 1999.

Variable	*Mean*	*Std. Dev.*
Workplace Practices		
Suggestion Program	0.311	0.463
Flexible Job Hours	0.326	0.469
Information Sharing	0.497	0.500
Teams	0.274	0.446
Committee	0.192	0.394
Workgroups	0.107	0.309
Cost of Absenteeism $E(w^a)$		
Group Incentives	0.150	0.357
Individual Incentives	0.421	0.494
Merit Pay	0.313	0.464
Profit-Sharing	0.159	0.365
Vacancy Rate	0.027	0.061
Layoff Rate	0.099	0.376
Size		
10–19 Employees	0.451	0.478
20–99 Employees	0.472	0.499
100–499 Employees	0.067	0.250
500 Employees and More	0.010	0.099
Number of Workplaces	3,767	

year for employees and at every fifth year for workplaces).

The initial sample comprises 23,540 employees in 1999, 20,167 of whom are also present in 2000. Workers are re-sampled from the same set of workplaces in 2001, yielding 20,352 individuals, of whom 16,813 are also present in 2002.

However, we exclude establishments with fewer than ten employees from the sample, because survey questions on work practices were not intended for them. Individuals who did not work throughout the year are included, but we control for their limited exposure to the risk of being absent in our regression framework. Finally, we drop workers whose absence resulted in their missing more than fifty days of work in the past year.[10]

The rich structure of the data set allows us to control for a variety of factors determining absenteeism decisions. From the worker questionnaire we are able to extract detailed demographic characteristics, including measures of health, human capital, and income from other sources. Moreover, we use detailed explanatory variables on the employment contract, including wage, contracted hours, and information about working hours flexibility and the exact scheduling of hours.

From the workplace questionnaire, we are able to construct firm size indicators and build measures of layoff and vacancy rates. Even more important, the establishment questionnaire includes very detailed information about current workplace practices (6). Finally, our regressions include industry (13), occupation (6), and time (4) dummies. Summary statistics on all explanatory variables are presented in Table 1 for the dependent variable, Table 2 for employees, and Table 3 for employers. Note that the number of weeks absent in Table 1 refers to a five-day workweek. Thus zero means the worker was absent fewer than five days during a year.

Results

Complete estimation results are presented in Table 4, where we contrast the determinants of absenteeism using only worker variables (column *I*), only job description variables (column *IJ*), only workplace variables (column *J*), and all variables (column *All*).[11] In all models, the dependent variable is the total number of days of absence that is reported for the whole year, including paid sick leave, other paid leave, and unpaid leave.[12] Using days of absence in this type of survey might be problematic if its distribution is not smooth. Moreover, it is possible that the respondent is not able to perfectly recall absences for a full year.[13] For these reasons, in our empirical analysis, we also tested the model using weeks of absence as the dependent variable. Since the results were unchanged, we show only results for days of absence.[14]

[10]Results are robust with respect to other cutoff points for eliminating outliers.

[11]Estimating the determinants of absence decisions using a Zero-inflated Poisson model yields similar results.

[12]"Other" paid leave does not include vacations, paternity/maternity leave, or absence due to strikes or lock-out.

[13]Unfortunately, it is not possible to compare the number of absences as reported by the worker with the number indicated by administrative measures.

[14]The structure of the data also does not allow us to study episodes of absenteeism.

Predictions of the theoretical model. In the first part of Table 4, we focus on the predictions of the theoretical model. We conclude that while most coefficients are of the expected sign, the magnitudes of the effects are rather small. We get a statistically significant negative coefficient for wages (w) in both specifications, although statistical significance drops to 10% in the complete specification. As expected, absences increased with contracted hours (t^c).[15] We also get the expected sign for non-labor income (R), although the effect is also quite small.

It is not clear what variables should be included in the empirical specification to proxy for the average cost of absenteeism ($E(w^a)$). In the literature, the cost of absenteeism is usually related to an increased likelihood of being fired or being passed up for promotion. Therefore, we settle on an indicator of the layoff rate (defined as the number of workers laid off in the past year divided by average employment) and the vacancy rate (defined as the number of positions available in the firm divided by average employment). These variables are interpreted as indicating the willingness of the workplace to use layoffs as a way to discipline employees. For example, a high vacancy rate might reflect an employer's reluctance to fire employees even if they misbehave. We also include measures of the use of incentive pay in the workplace. The absent worker might be compensated for lost wages due to absence, but it is conceivable that the probability of receiving merit pay, a share of the profits, or group incentives will diminish as a result of his absence.

We find that the layoff rate had a negligible impact on absences. It is possible that workplaces with high layoff rates were moving toward bankruptcy, in which case absence is not likely to have been costly anyway. However, as expected, high vacancy rates were associated with higher absences. While we interpret this as a direct impact of a low cost of absenteeism, an alternative explanation for the sign and magnitude of the coefficient focuses on the fact that workers in such workplaces are more likely to overwork and rely on absences to relieve the physical and mental pressures of this additional work. Among incentive-pay variables, only the coefficient on the use of merit pay is statistically significant at the 1% level, and it is not of the expected sign. It should be noted that this variable indicates whether workplaces used merit pay and not whether the particular employee could receive merit pay. If merit pay was available only to management employees, a result could be lower work force morale and increased absence rates.

Demographics, health, and human capital. We find that women were more likely than men to be absent. To be precise, women had 1.3 times the absence rate of men, but being married reduced absenteeism. Contrary to some previous studies,[16] we do not find that women with children had higher levels of absenteeism; perhaps childcare is a more equally shared responsibility now than in the past. Health is also found to have been a very important determinant of the absenteeism decision. The absence rate for individuals who had no activity limitation was 25% below that for individuals with some limitation. Like some previous researchers, we find that ethnicity was not related to absence.

Education was positively associated with absences, but the impact peaked for workers with a bachelor's degree. Seniority was also positively associated with absences, an outcome that might reasonably be expected if greater seniority is linked with greater job security (Barmby, Ercolani, and Treble 2002). Experience (or age) was related to lower absence, but at an increasing rate. This pattern is usually explained by the hypothesis that older workers expect longer unemployment spells if fired and are therefore careful not to engage in behavior that might lead to job loss.

[15] This is in contrast to Allen (1981) and Vistnes (1997), for example. We note that it is important to take into account unobserved heterogeneity in this case, since individuals working longer hours might enjoy their work more (and thus be absent less frequently).

[16] Vistnes (1997) found a statistically significant interaction between being a woman and having young children.

Table 4. Count Model with Unobserved Heterogeneity on Days of Absence. (Robust Standard Errors in Parentheses)

Variable	*I*	*IJ*	*J*	*All*
Variables from the Theoretical Model				
Natural Log of Hourly Wage (w)		–0.055***		–0.011*
		(0.004)		(0.006)
Contracted Hours (t^c)		0.000		0.002***
		(0.000)		(0.000)
Income from Other Sources (000s) (R)	0.000***			0.000***
	(0.000)			(0.000)
Cost of Absenteeism ($E(w^a)$):				
Group Incentives		0.005		0.004
		(0.006)		(0.016)
Individual Incentives		0.012**		0.022*
		(0.005)		(0.013)
Merit Pay		0.019***		0.048***
		(0.005)		(0.012)
Profit-Sharing		0.090***		0.011
		(0.006)		(0.016)
Vacancy Rate			0.071***	0.085***
			(0.024)	(0.030)
Layoff Rate			0.008***	0.009***
			(0.002)	(0.002)
Demographic Characteristics				
Women	0.256***			0.250***
	(0.010)			(0.010)
Black	–0.025			–0.067
	(0.036)			(0.041)
Other Race	–0.087***			–0.039***
	(0.010)			(0.010)
Married	–0.023***			–0.034***
	(0.005)			(0.005)
Number of Pre-School-Aged Kids	0.004			0.007
	(0.005)			(0.005)
Women × Pre-School-Aged Kids	–0.002			0.001
	(0.008)			(0.008)
Health				
No Activity Limitation	–0.260***			–0.257***
	(0.004)			(0.004)
Human Capital				
High School Degree	–0.002			0.030***
	(0.011)			(0.012)
Less Than Bachelor Degree	0.161***			0.139***
	(0.011)			(0.011)
Bachelor Degree	0.119***			0.109***
	(0.015)			(0.016)
Some Higher Education	0.034*			0.062***
	(0.019)			(0.020)
Seniority	0.046***			0.043***
	(0.001)			(0.001)
Seniority Squared (/100)	–0.122***			–0.117***
	(0.004)			(0.004)
Experience	–0.010***			–0.011***
	(0.001)			(0.001)
Experience Squared (/100)	0.009***			0.011***
	(0.003)			(0.003)
Work Arrangement				
Work Regular Hours		–0.065***		–0.038***
		(0.003)		(0.004)

Continued

Table 4. Continued.

Variable	*I*	*IJ*	*J*	*All*
Work on Weekend		–0.031***		–0.032***
		(0.004)		(0.004)
Work Flexible Hours		0.037***		0.023***
		(0.002)		(0.003)
Work Nontraditional Working Hours		0.038***		–0.002
		(0.003)		(0.004)
Work at Home		–0.030***		–0.019***
		(0.003)		(0.004)
Work in Shift		0.010**		0.024***
		(0.004)		(0.005)
Work a Reduced Workweek		–0.080***		–0.048***
		(0.005)		(0.006)
Work a Compressed Workweek		0.094***		0.113***
		(0.005)		(0.006)
Collective Bargaining Agreement		0.228***		0.244***
		(0.004)		(0.006)
Technology				
Use a Computer		0.059***		0.013***
		(0.003)		(0.004)
Use Computer-Assisted Design		0.030***		0.025***
		(0.003)		(0.004)
Use Other Technology		0.111***		0.096***
		(0.002)		(0.003)
Workplace Practices				
Employee Suggestions			0.110***	0.050***
			(0.005)	(0.012)
Flexible Job Design			–0.060***	–0.007
			(0.005)	(0.014)
Information Sharing with Employees			0.037***	0.069***
			(0.005)	(0.013)
Problem-Solving Teams			–0.042***	–0.016
			(0.005)	(0.013)
Labor-Management Committee			0.207***	0.070***
			(0.005)	(0.013)
Self-Directed Workgroups			0.010*	–0.010
			(0.006)	(0.014)
Firm Size				
20–99 Employees			0.103***	0.059***
			(0.007)	(0.007)
100–499 Employees			0.171***	0.114***
			(0.007)	(0.009)
500+ Employees			0.206***	0.175***
			(0.008)	(0.011)
Other Parameters				
Constant	–0.011	0.297***	–0.095***	–0.302***
	(0.038)	(0.022)	(0.016)	(0.045)
σ_θ	1.291***	0.977***	0.982***	1.282***
	(0.004)	(0.002)	(0.002)	(0.004)
σ_ψ	0.402***	0.376***	0.378***	0.429***
	(0.007)	(0.002)	(0.002)	(0.007)
Ln-Likelihood	–183,270.03	–188,546.61	–188,748.51	–182,890.63
Observations	64,367	64,367	64,367	64,367

Note: Industry (13), occupation (5), region (6) and year (3) dummies are included in all regressions.
*Statistically significant at the .10 level; **at the .05 level; ***at the .01 level.

Work arrangement and technology. It has been said that new work arrangements lead to more stress and more absenteeism. While we find this to be partly true—in particular, we find that workers working regular hours were less often absent than their counterparts—our results strongly suggest that it was the case for certain work arrangements only. Using detailed data on the scheduling of the work week, we find that workers who worked at home or worked on a reduced work week had a lower incidence of absenteeism, while workers with flexible hours, shift work, or assignment to a compressed workweek had higher levels.[17] Contrary to prior expectations, we find that working weekend hours or other non-traditional hours did not increase absences.

Finally, we find that workers covered by a collective bargaining agreement were more often absent than were non-covered workers. This positive relationship between unionization and absences has been explained in the literature by the fact that collective bargaining agreements are often accompanied by a grievance process that could make it difficult for the workplace to fire workers with frequent absences (Vistnes 1997). Technology use was associated with higher absence, although the effect was much stronger for workers using technologies associated with repetitive work (see the coefficient for the variable "other technologies," which includes cash registers, sales terminals, scanners, and so on).

Workplace practices and firm size. Firms would normally be interested in finding what organizational practices succeed in reducing absenteeism when it is costly. In the regression with only firm-level variables (column *J*), we find that the use of flexible job design and problem-solving teams was associated with lower levels of absenteeism. It has been noted in the literature that the use of teams increases the cost of absenteeism (Heywood and Jirjahn 2004), so these results are not unexpected. However, once we control for other worker and job characteristics, those effects are no longer statistically significant, underlining the importance of correctly specifying the set of regressors. We also find that employee suggestion programs, information sharing, and labor-management committees were linked to more absences. This result might be explained by reverse causality if these particular workplace practices are introduced partly in response to high levels of absenteeism in the workplace. Wilson and Peel (1991), who obtained a similar result with respect to worker participation in decision-making, speculated that a more relaxed approach to absenteeism might be adopted in these workplaces. However, they also found that firms with participation schemes had significantly lower average absenteeism than other firms, a finding that is not corroborated by our data.

It should be noted that we find statistically significant workplace size effects even though we control for schedule flexibility (see Vistnes 1997). This means that lower absenteeism in smaller workplaces was probably due not to more flexible work arrangement but, more likely, to higher employee attachment or involvement levels in these workplaces.

The impact of using linked data. We would first point out that both unobserved heterogeneity components are statistically significant. It is interesting to note that intra-workplace unexplained variation in absences is more important than inter-workplace unexplained variation. Basically the same pattern is found for observed characteristics: the likelihood levels show that worker characteristics have greater explanatory power than either job or workplace characteristics.

Comparing the different columns of Table 4, it should be noted that omitting either individual (*I*), job (*IJ*), or workplace (*J*) characteristics leads to notable biases in the magnitude of the estimated coefficients. For example, the estimated wage effect becomes smaller in the full specification and less statistically significant. This means that wage increases are unlikely to reduce absences appreciably. As noted above, there are also important biases in the estimated impact of workplace practices. Also, the impact of using a computer is greatly overestimated in the

[17]Thus, in contrast to Allen (1981), we do not find higher absence rates for workers who have the same schedule every week or who work standard hours.

specification using only job characteristics. The same can be said for firm size effects in the specification that uses only workplace characteristics. One exception to this general rule, however, is the relative stability of estimates for demographic and health characteristics across specifications.

Conclusion

We have examined the factors associated with the absenteeism decision at the worker level. We draw on a survey that is unusual in linking information on the worker with information on the workplace, a feature that has allowed us to control for detailed demographic, worker, job, and firm characteristics, as well as unobserved heterogeneity at both the worker and workplace levels.

Overall, we find that workplaces had considerable leverage in reducing absenteeism. Our results show that success in this objective was more likely to come through the choice of appropriate work arrangements than through wage increases. For example, we find strong evidence that standard working hours (Monday–Friday between 6:00 a.m. and 6:00 p.m.), work-at-home options, and reduced workweeks were associated with lower absences. The reverse was true for shift work and compressed work weeks. We also find that in the complete specification, workplace practices such as problem-solving teams and flexible job design had no impact, a result that underscores the importance of correctly specifying the set of regressors.

The use of incentive pay in the workplace is shown to have had ambiguous effects on absences in our data. More detailed data on incentive pay would be useful to sort out the effects. Finally, it would be interesting to see if the determinants of absenteeism differ depending on the type of absenteeism (paid/unpaid leave, sick leave, other leave).

REFERENCES

Abowd, John M., and Francis Kramarz. 1999. "The Analysis of Labor Markets Using Matched Employer-Employee Data." In Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics,* Vol. 3B, Chap. 40, pp. 2629–2710. Amsterdam: Elsevier Science (North Holland).

Akyeampong, Ernest B. 2005. "Fact Sheet on Work Absences." *Perspectives on Labour and Income,* Vol. 6, No. 4, pp. 21–30.

Allen, Steven G. 1981. "An Empirical Model of Work Attendance." *Review of Economics and Statistics,* Vol. 63, No. 1, pp. 77–87.

____. 1983. "How Much Does Absenteeism Cost?" *Journal of Human Resources,* Vol. 18, No. 3, pp. 379–93.

Barmby, Tim. 2002. "Worker Absenteeism: A Discrete Hazard Model with Bivariate Heterogeneity." *Labour Economics,* Vol. 9, No. 4, pp. 469–76.

Barmby, Tim, Marco Ercolani, and John Treble. 2002. "Sickness Absence: An International Comparison." *Economic Journal,* Vol. 112, No. 480, pp. F315–F331.

Barmby, Tim, Chris Orme, and John G. Treble. 1991. "Worker Absenteeism: An Analysis Using Microdata." *Economic Journal,* Vol. 101, No. 405, pp. 214–29.

Blomquist, N. Sören. 1983. "The Effect of Income Taxation on the Labor Supply of Married Men in Sweden." *Journal of Public Economics,* Vol. 22, No. 2, pp. 169–97.

Delgado, Miguel A., and Thomas J. Kniesner. 1997. "Count Data Models with Variance of Unknown Form: An Application to a Hedonic Model of Worker Absenteeism." *Review of Economics and Statistics,* Vol. 79, No. 1, pp. 41–49.

Dionne, Georges, and Louis Eeckhoudt. 1987. "Proportional Risk Aversion, Taxation, and Labour Supply under Uncertainty." *Journal of Economics,* Vol. 47, No. 4, pp. 353–66.

Drago, Robert, and Mark Wooden. 1992. "The Determinants of Labor Absence: Economic Factors and Workgroup Norms across Countries." *Industrial and Labor Relations Review,* Vol. 45, No. 4, pp. 764–78.

Dunn, Lucia F., and Stuart A. Youngblood. 1986. "Absenteeism as a Mechanism for Approaching an Optimal Labor Market Equilibrium: An Empirical Study." *Review of Economics and Statistics,* Vol. 68, No. 4, pp. 668–74.

Ehrenberg, Ronald G. 1970. "Absenteeism and the Overtime Decision." *American Economic Review,* Vol. 60, No. 3, pp. 352–57.

Frankel, Emil. 1921. "Labor Absenteeism." *Journal of Political Economy,* Vol. 29, No. 6, pp. 487–99.

Gilleskie, Donna B. 1998. "A Dynamic Stochastic Model of Medical Care Use and Work Absence." *Econometrica,* Vol. 66, No. 1, pp. 1–45.

Hausman, Jerry A. 1980. "The Effect of Wages, Taxes, and Fixed Costs on Women's Labor Force Participation." *Journal of Public Economics,* Vol. 14, No. 2, pp. 161–94.

Henrekson, Magnus, and Mats Persson. 2004. "The Effects on Sick Leave of Changes in the Sickness Insurance System." *Journal of Labor Economics,* Vol. 22, No. 1, pp. 87–113.

Heywood, John S., and Uwe Jirjahn. 2004. "Teams, Teamwork, and Absence." *Scandinavian Journal of Economics,* Vol. 106, No. 4, pp. 765–82.

Johansson, Per, and Marten Palme. 1996. "Do Economic Incentives Affect Work Absence? Empirical Evidence Using Swedish Micro Data." *Journal of Public Economics,* Vol. 59, No. 2, pp. 195–218.

____. 2002. "Assessing the Effect of Public Policy on Worker Absenteeism." *Journal of Human Resources,* Vol. 37, No. 2, pp. 381–409.

Kauermann, Göran, and Renate Ortlieb. 2004. "Temporal Pattern in Number of Staff on Sick Leave: The Effect of Downsizing." *Journal of the Royal Statistical Society, Series C-Applied Statistics,* Vol. 53, pp. 355–67.

Kenyon, Peter, and Peter Dawkins. 1989. "A Time Series Analysis of Labour Absence in Australia." *Review of Economics and Statistics,* Vol. 71, No. 2, pp. 232–39.

Rothschild, Michael, and Joseph Stiglitz. 1971. "Increasing Risk II: The Economic Consequences." *Journal of Economic Theory,* Vol. 3, pp. 66–84.

Vistnes, Jessica P. 1997. "Gender Differences in Days Lost from Work Due to Illness." *Industrial and Labor Relations Review,* Vol. 50, No. 2, pp. 304–23.

Wilson, Nicholas, and Michael J. Peel. 1991. "The Impact on Absenteeism and Quits of Profit-Sharing and Other Forms of Employee Participation." *Industrial and Labor Relations Review,* Vol. 44, No. 3, pp. 454–68.