Coughs and Sneezes Spread Diseases: An Empirical Study of Absenteeism and Infectious Illness

Tim Barmby[1] (Department of Economics, University of Aberdeen)

Makram Larguem (ERMES, CNRS, Université de Paris 2, Panthéon-Assas)

Nov 2007

Abstract: This paper incorporates some theoretical ideas from the study of the epidemiology of infectious illness into a model of worker absence. The paper then seeks to quantify such infection effects by examining a personnel dataset which allows us to track daily absence decisions of a group of industrial workers employed in the same factory. We find significant effects of our measure of sickness prevalence in the (rest of the) workforce on the absence probabilities of individual workers.

Keywords: Absence, Infectious illness

JEL Codes J33, I12

1. Introduction

A number of recent studies have suggested that workers' attendance as well as their absence behaviour, could have implications for the way in which firms' design remuneration contracts. The reason for this is that, since worker absenteeism is in large part due to illness, if contractual requirements are sufficiently stringent to induce workers to attend work when ill this could result in the illness being more readily communicated to other workers with associated effects on productivity. The papers by Chatterji and Tilley (2000) and Skåtun (2002) develop the theoretical framework for

[1] We acknowledge helpful comments from John Skåtun , John Treble, Tor Eriksson, Ali Skalli, Mike Nolan and Colin Tilley. We, of course, are responsible for remaining errors. Corresponding author; Tim Barmby, Department of Economics, University of Aberdeen, Old Aberdeen, Scotland AB24 3QY. tim.barmby@abdn.ac.uk

this idea, this paper seeks to quantify the effect of the potential for infection by examining a personnel dataset which allows us to track daily absence decisions of a group of industrial workers employed in the same factory.

2. Models of Absenteeism

The bulk of existing work on absenteeism, dating from an early study by Buzzard and Shaw (1952), Allen (1981), Dunn and Youngblood (1986) and many others have taken, either implicitly or explicitly, a labour supply perspective on observed worker absence. In this labour supply framework firms set remuneration contracts (which in their most basic form will comprise an hourly wage rate and a level of contracted hours) and workers respond, firstly by deciding whether to accept the contract out of those on offer and then whether to deviate from the contract by supplying *less* than the contracted hours specified. In this sense then absence can only be defined if the contract specifies contracted hours.

Only recently have some researchers, most notably Chatterji and Tilley, and Skåtun, in the papers mentioned above, sought to explore further the impact of remuneration contracts. In particular, if the contract induces workers who are ill to attend work, might this result in higher subsequent general absence. In this section we will try and incorporate some theoretical ideas from the study of infectious illness and epidemiology into an empirical model of worker absence, being guided by the framework used by Philipson (2000) and Skåtun (2003).

To do this we first consider the nature of the data we are using, (we describe more specific details of the data in terms of covariates in the next section, the purpose of introducing data consideration at this point is to identify how we need to formulate the epidemiological links). The data we have are event histories consisting of sequences of realisations of indicator random variable d_{it} which equals 1 if the worker is absent on a given day and zero otherwise, the subscript i indicates the worker and t the time period, (workers in general will be observed for different lengths of time

T_i[2]). The utility which a worker attaches to each of the two available choices is described by

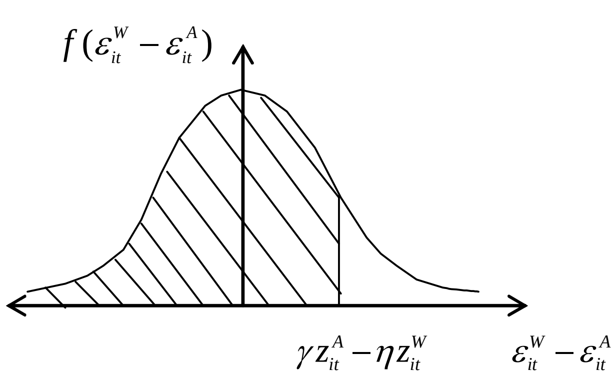
$$\begin{aligned} U_{it}^{W} &= \eta z_{it}^{W} + \varepsilon_{it}^{W} \\ U_{it}^{A} &= \gamma z_{it}^{A} + \varepsilon_{it}^{A} \end{aligned} \quad i = 1,....,N; t = 1,.......,T_i \tag{1}$$

Where η and γ are conformable (row) vectors describing the effect on utility of data for both characteristics of alternatives and of the worker contained in z. Data on alternatives here would be income in each of the two states, so the utility which an individual would attach to absence would be related to their sickpay and the utility they would attach to work would be related to their (expected) earnings. Data on the characteristics on individuals, gender etc will of course not vary over alternatives, but can be thought of as influencing the utility of work and absence differently. The probability of the event of observing worker *i* absent at time *t* can be written

$$P(d_{it} = 1) = P(U_{it}^{A} > U_{it}^{W}) = P(\varepsilon_{it}^{W} - \varepsilon_{it}^{A} < \gamma z_{it}^{A} - \eta z_{it}^{W}) = F(\beta x_{it}) \tag{2}$$

where F is the CDF of $\varepsilon_{it}^{W} - \varepsilon_{it}^{A}$, and for data which varies between alternatives $\gamma = \eta = \beta$, and for characteristics or individuals which don't vary across alternatives β is interpreted as $\beta = (\gamma - \eta)$. It will be useful to dwell a little on the way in which the model structure represents the occurrence of illness, and how it effects the actions of workers.

Figure 1: The interpretation of $\varepsilon_{it}^{W} - \varepsilon_{it}^{A}$ as worker's health state



[2] In this paper we don't explicitly model the potential endogeneity of T_i .

The hatched area is the event defined in (2), that the worker is absent. Since, the vector z_{it}^{A} could contain as one of its elements the level of sickpay, and that since the corresponding coefficient on this variable would, given the utility interpretation in equation (1), be positive, would suggest that higher sickpay would increase the probability of absence. Similarly since z_{it}^{W} could contain the worker's wage rate, wage should have a negative effect, as is commonly found.

The worker's health state is represented in this model by random error terms ε_{it}^{W} and ε_{it}^{A}; when a worker is ill he/she attaches more utility to not working, it is not identified whether this means higher ε_{it}^{A} or lower ε_{it}^{W} just that $\varepsilon_{it}^{W} - \varepsilon_{it}^{A}$ is lower. A lower $\varepsilon_{it}^{W} - \varepsilon_{it}^{A}$ is consistent with higher absence as is indicated in figure 1.

It is clear from the above formulation that the probability of absence is driven by the difference in the systematic component of the equations in (1) $\gamma z_{it}^{A} - \eta z_{it}^{W}$, if the z_{it}^{A} vector contains sickpay and the z_{it}^{W} vector earnings then the probability of absence should be a function of the difference. This is reflected in the way we specify the probability in our empirical analysis.

The central empirical problem faced by this paper is to specify the above probability in such a way that we can interpret the estimated parameters in terms of a coherent model of worker illness. To this end we first consider the event that an absence spell *starts* at time t, this is the event $(d_{i,t} = 1 | d_{i,t-1} = 0)$. At t-1 a worker *i* was *exposed* to $N - \sum_{\substack{j=1 \\ j \neq i}}^{N} d_{j,t-1}$ workers who were potentially infected. The proportion of these who are infected is in epidemiological terminology the *prevalence*, which we denote, *s*. The *transmission* rate of the infection, μ, describes the ease with which the sickness can be passed from an infected to a non-infected individual, see Philipson (2000). The probability of a healthy worker being infected, θ, can therefore be written.

$$\theta \propto \mu s(N - \sum_{\substack{j=1 \\ j \neq i}}^{N} d_{j,t-1}) \tag{3}$$

An estimate of s, the prevalence of sickness, can, we argue, be formed by observing how many workers (apart from i) who attended work at t-1 but were absent at t (this is the *number* of newly started spells at t), and taking this as a ratio to the total number at work at t-1.

$$\tilde{s} = \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} (d_{j,t} \mid d_{j,t-1} = 0)}{N - \sum_{\substack{j=1 \\ j \neq i}}^{N} d_{j,t-1}} \tag{4}$$

this implies that an estimate of the probability of infection can be written

$$\tilde{\theta} \propto \mu \frac{\sum_{\substack{j=1 \\ j \neq i}}^{N} (d_{j,t} \mid d_{j,t-1} = 0)}{N - \sum_{\substack{j=1 \\ j \neq i}}^{N} d_{j,t-1}} (N - \sum_{\substack{j=1 \\ j \neq i}}^{N} d_{j,t-1}) = \mu \sum_{\substack{j=1 \\ j \neq i}}^{N} (d_{j,t} \mid d_{j,t-1} = 0) \tag{5}$$

Note that to be infected the worker has to be at work, also if the worker is absent then he/she is not exposed to infection. We assume, therefore, that the probability of absence will be an (increasing) function of θ multiplied by $(1-d_{i,t-1})$. These, taken together, suggest the following form for the probability in equation (2).

$$P(d_{it}=1) = F(\mu(1-d_{i,t-1}) \sum_{\substack{j=1 \\ j \neq i}}^{N} (d_{j,t} \mid d_{j,t-1} = 0)) \tag{6}$$

This has a simple interpretation; the probability of absence for a working individual *i* is a function of the number of *newly* started absence spells amongst the rest of the labour force, and that the coefficient will be an estimate of the transmission rate μ . Essentially what this method is picking up is whether there exists any correlation between the *incidence* of absence spells starting. Since this suggested empirical method is the new contribution of our paper we illustrate in detail the construction of the regressor in (6) in appendix 1, for a constructed example.

The term in the regression component here is, of course, also only capturing the potential effect of our model of the worker illness, there will be other factors which we must take into account. To fully specify the probability function we also include

1) x_{it} : these are covariates, either time variant or, invariant characteristics of workers, here we include days of the week, gender and contracted hours

2) $d_{i,t-1}$: this captures the effect of *own* health state. If the worker was absent last period how does this affect the probability of being absent this period. Heckman (1981) terms this structural dependence

3) $\sum_{s=1}^{t-1}\prod_{k=1}^{s} d_{i,t-1}$: this captures the effect of duration in the state, this can be entered in a as a polynomial to capture potential non-linearities.

Adding these terms gives us an extended probability model of the form

$$P(d_{it}=1)=\underbrace{F(\beta' x_{it}+\gamma d_{i,t-1}+\lambda\sum_{s=1}^{t-1}\prod_{k=1}^{s} d_{i,t-1}+\mu(1-d_{i,t-1})\sum_{\substack{j=1\\ j\neq i}}^{N}(d_{j,t}\mid d_{j,t-1}=0)+\sigma u_i)}_{\theta Z}$$

$$i=1,....,N \quad t=1,......,T_i \quad \theta' =(\beta\ \gamma\ \lambda\ \mu)$$

$$Z' =[x_{it}, d_{i,t-1}, \sum_{s=1}^{t-1}\prod_{k=1}^{s} d_{i,t-1}, (1-d_{i,t-1})\sum_{\substack{j=1\\ j\neq i}}^{N}(d_{j,t}\mid d_{j,t-1}=0)] \qquad (7)$$

We include an unobserved term u_i in (7) since the data we are working with is gathered from a firm's personnel records. As such it will primarily include data which it needs for the accurate operation of its payroll, this means that certain data which could affect the probability of absence (whether the household has young children for instance) is not observed. So we will have to take the existence of unobserved heterogeneity seriously. We assume the unobserved term in (7) is a drawing from N(0,1) density $\phi(u)$ describing the distribution of the unobserved term in the population from which the sample is drawn, given this assumption we can integrate out to form a marginal likelihood,

$$LnL(\theta,\sigma) = \sum_{i=1}^{N} Ln \int_{-\infty}^{\infty} \prod_{t=2}^{T_i} F(\theta Z_{it} + u)^{d_{it}} (1 - F(\theta Z_{it} + u))^{(1-d_{it})} \phi(u) du \quad (8)$$

the CDF F is assumed Logistic[3].

3. Data

Our data is a sample of 944 workers in 1988 drawn from a UK manufacturing firm which produces a homogeneous product using production lines. We have data from the personnel records giving us their absence for days they were contracted to work on and some individual characteristics, gender and marital status which we include in the specification. The main aspects of the remuneration contract for which we have information on are wage rates, contracted hours and the sick pay for which each worker is eligible. The wage information is in the form of weekly payments recorded in the payroll. A worker's remuneration is made up of a basic pay component, overtime payments and shift premia. Basic pay represents the bulk of overall earnings.

The firm operates a sick pay scheme which defines the eligibility of workers to sick pay in excess of a minimum Statutory Sick Pay (SSP). This eligibility is based on whether their past attendance and so is an *experience rated* system. It is the operation of this sick pay scheme, which is the main way in which the cost of a day's

[3] This corresponds to $\varepsilon_{it}^{W}, \varepsilon_{it}^{A}$ in (1) being IID Extreme Value Type I

absence varies across individuals. A worker's attendance record is measured by a points scheme where each day of certain types of absence attracts one point[4]. The categories of absence which attract points are, unexplained absence, self certified absence (in the UK workers can self certify themselves absent for up to one week before having to obtain a medical certificate), and medically certified absence.

Workers are eligible for sick pay at one of three levels of generosity, A, B and C, depending on their points total over a rolling two year period[5]. Workers were regraded at the end of each year. As can be seen the difference between the sickpay the worker was entitled to and his normal earnings was determined primarily by his/her sickpay grade. We take account of this in the specification of our model.

For our data the mean absence rate is less than 3% (2.73) so for a workforce of 944, this will mean that around 26 workers will be absent on a given day but since the duration of these workers spells of absence will vary, the number of newly started spells of absence on a given day will be smaller, this is reported in appendix 2

4. Results

The results reported in Table 1 are supportive of the notion that contract effects are important in explaining absence patterns. The gender dummy (female = 1) has a positive significant coefficient. Normal daily wage has a negative effect and when interacted with whether the individual is in sick-pay grades B and C where there is the highest cost of a days absence to the worker this becomes more negative, these results are similar to Barmby (2002)[6]. There is positive duration dependence peaking around 4 days. The increasingly negative coefficient on day of the week indicates that (ceteris paribus) absence rates fall during the week.

---

[4] during this period UK firms were able to recover some sickpay costs from the government if the absence spell was not preceded by another recent spell. If this was the case then the firm would attach 2 points not 1 to each day of the spell.

[5] If the average over the previous 2 years was less than 10 points per year then the worker was entitled to essentially full replacement of normal earnings (this is A grade), B grade was between 10 and 20 and workers in B grade were only allowed sick pay equal to their basic component of earnings. Workers with more than 20 points on average over the previous 2 years were graded C and receive no company sick pay

[6] Note the method of interacting a regressor with $d_{i,t-1}$ allows the coefficient to vary between states (or work and absence) as discussed in Barmby (1998)

In the framework we have set up there are two main sources of variation in illness (and therefore possible absence). The first is through own health state, which we represent in terms of the structural dependence on lag of own absence $d_{i,t-1}$. Secondly through the influence of the estimated prevalence of illness of other workers, which is measured by the number of newly started spells variable, NSS, the construction of which we discussed earlier.

The lag of own absence $d_{i,t-1}$ gives a significant positive effect, as intuition would suggest. Since we have in mind that absence is being generated by some underlying stochastic illness process as we discussed earlier, we have to take into account the influence of the weekend.

Table 1 Random Effects estimation of absence model.N= 944, NT = 181604, mean absence rate 0.0273

Variable	Coeff.	Std. Err
Worker characteristics		
Gender (female=1)	0.291	(0.073)***
Contract effects		
Normal wage	-0.002	(0.001)***
Wage*grade B	-0.002	(0.0014)
Wage*grade C	-0.003	(0.0019)*
Contracted hours	0.029	(0.001)***
Contract hours* $d_{i,t-1}$	-0.020	(0.010)**
Grade B	0.442	(0.166)***
Grade C	0.645	(0.216)***
Absence effects		
Lagged absence ($d_{i,t-1}$)	6.162	(0.446)***
Duration of absence	1.128	(0.223)***
Duration² of absence	-0.141	(0.047)***
(1- $d_{i,t-1}$)*NSS	0.122	(0.006)***
Day effects		
Tuesday	-7.26	(0.383)***
Wednesday	-7.62	(0.384)***
Thursday	-7.72	(0.384)***
Friday	-8.13	(0.391)***
σ	0.457	(0.044)***
Log likelihood	-6325.14	

The significance of the estimated coefficients are given by: * (10%), **(5%) and ***(1%). Note also that the day dummies Tuesday through Friday negate the need for a constant, since there are no Mondays in the sample, this comes from the construction of (1- $d_{i,t-1}$)*NSS as appendix 1 makes clear

5. Concluding Remarks

The model developed above shows how an individual workers propensity to absent themselves from work is affected by a range of factors including their personal characteristics and the terms of their remuneration contract. We find that absence is positively affected by a worker's own absence, which is in line with other work. We extend previous work to construct a model of absence incorporating an epidemiological structure and find significant effects of our measure of sickness prevalence in the (rest of the) workforce on the absence probabilities of individual workers. This provides empirical support for both common-sense intuition and theoretical models exploring further aspects of the way in which worker illness affect the way in which firms might set their working arrangements.

References

Allen S (1981) "An Empirical Model of Work Attendance" *The Review of Economics and Statistics* Vol 63 pp 77-87.

Barmby T (2002) "Worker Absenteeism: A Discrete Hazard Model with Bivariate Heterogeneity" *Labour Economics* Vol 9 pp 469-476.

Barmby T (1998) "The Relationship between Event History and Discrete Time Duration Models: an Application to the analysis of Personnel Absenteeism" *Oxford Bulletin of Economics and Statistics* Vol 60(2) pp 261-265

Buzzard R B and Shaw W J (1952) "An Analysis of Absence Under a Scheme of Paid Sick Leave" *British Journal of Industrial Medicine* Vol 9 pp 282-295

Chatterji M and Tilley C (2002) "Sickness Absenteeism, Presenteeism, and Sickpay" *Oxford Economic Papers* Vol 54 pp 669-87

Dunn L.F and Youngblood S A (1986) "Absenteeism as a Mechanism for approaching an Optimal Labor Market Equilibrium" *The Review of Economics and Statistics* Vol 68 pp 668-73

Heckman J.J. (1981) "Statistical models for discrete panel data" In Structural Analysis of Discrete Data with Econometric Applications (ed C.F.Manski and D.McFadden) Cambridge, Mass. MIT press.

Philipson T (2000) "Economic Epidemiology and Infectious Diseases" Chapter 33 in Handbook of Health Economics Vol 1B by Culyer A J and Newhouse J P.

Skåtun J (2003) "Take some time off, why don't you : endogenous sickleave and pay", *Journal of Health Economics*. Vol. 22, issue 3, Pages 379-402

Appendix 1

The construction of the variable “newly started spells” NSS = $\sum_{\substack{j=1\\ j\neq i}}^{N}(d_{j,t}\,|\,d_{j,t-1}=0)$ is

$$absence\left\{\overbrace{\begin{matrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{matrix}}^{Week}\right.$$

$$NSS\left\{\begin{matrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 & 0 \\ 1 & 2 & 0 & 2 & 0 \\ 1 & 1 & 0 & 2 & 0 \\ 1 & 2 & 0 & 1 & 0 \end{matrix}\right.$$

$$(1-d_{i,t-1})NSS\left\{\begin{matrix} - & 1 & 0 & 1 & 0 \\ - & 0 & 0 & 0 & 0 \\ - & 2 & 0 & 2 & 0 \\ - & 1 & 0 & 0 & 0 \\ - & 2 & 0 & 1 & 0 \end{matrix}\right.$$

“-“ meaning not defined. With weekends the situation looks like

$$
absence\left\{\begin{matrix}
\overbrace{\begin{matrix}0 & 1 & 0 & 1 & 0\end{matrix}}^{Week} & \overbrace{\begin{matrix}- & -\end{matrix}}^{Weekend} & \overbrace{\begin{matrix}0 & 1 & 0 & 1 & 0\end{matrix}}^{Week} \\
\begin{matrix}1 & 1 & 1 & 1 & 1\end{matrix} & \begin{matrix}- & -\end{matrix} & \begin{matrix}1 & 1 & 1 & 1 & 1\end{matrix} \\
\begin{matrix}0 & 0 & 0 & 0 & 0\end{matrix} & \begin{matrix}- & -\end{matrix} & \begin{matrix}0 & 0 & 0 & 0 & 0\end{matrix} \\
\begin{matrix}0 & 1 & 1 & 0 & 0\end{matrix} & \begin{matrix}- & -\end{matrix} & \begin{matrix}0 & 1 & 1 & 0 & 0\end{matrix} \\
\begin{matrix}0 & 0 & 0 & 1 & 1\end{matrix} & \begin{matrix}- & -\end{matrix} & \begin{matrix}0 & 0 & 0 & 1 & 1\end{matrix}
\end{matrix}\right.
$$

$$
NSS\left\{\begin{matrix}
1 & 1 & 0 & 1 & 0 & - & - & 1 & 1 & 0 & 1 & 0 \\
0 & 2 & 0 & 2 & 0 & - & - & 0 & 2 & 0 & 2 & 0 \\
1 & 2 & 0 & 2 & 0 & - & - & 1 & 2 & 0 & 2 & 0 \\
1 & 1 & 0 & 2 & 0 & - & - & 1 & 1 & 0 & 2 & 0 \\
1 & 2 & 0 & 1 & 0 & - & - & 1 & 2 & 0 & 1 & 0
\end{matrix}\right.
$$

$$
(1-d_{i,t-1})NSS\left\{\begin{matrix}
- & 1 & 0 & 1 & 0 & - & - & - & 1 & 0 & 1 & 0 \\
- & 0 & 0 & 0 & 0 & - & - & - & 0 & 0 & 0 & 0 \\
- & 2 & 0 & 2 & 0 & - & - & - & 2 & 0 & 2 & 0 \\
- & 1 & 0 & 0 & 0 & - & - & - & 1 & 0 & 0 & 0 \\
- & 2 & 0 & 1 & 0 & - & - & - & 2 & 0 & 1 & 0
\end{matrix}\right.
$$

Appendix 2

Table A1: Distribution of number of newly started spells (NSS) of absence

NSS	Frequency.	Percentage	Cum.
0	73,973	40.73	40.73
1	52,345	28.82	69.56
2	18,960	10.44	80.00
3	10,647	5.86	85.86
4	8,490	4.68	90.53
5	6,411	3.53	94.07
6	2,145	1.18	95.25
7	766	0.42	95.67
8	1,322	0.73	96.40
9	746	0.41	96.81
10	2,237	1.23	98.04
12	13	0.01	98.05
13	670	0.37	98.41
14	1,422	0.78	99.20
15	16	0.01	99.21
16	760	0.42	99.63
27	28	0.02	99.64
28	653	0.36	100.00
Total	181,604	100.00	