Anett C. Hansen and Harald K. Selte

Air Pollution and Sick-leaves – is there a Connection? A Case Study using Air Pollution Data from Oslo

Abstract:
In recent years a growing number of studies have been discussing the relationship between air pollution and human health. The evidence in the literature for adverse health effects of several pollutants seems convincing. In our article we are concerned with to which extent these health effects in turn induce sick-leaves or other kinds of reduced labour productivity, which is important for assessment of air pollution costs. We analyse the association between sick-leaves in a large office in Oslo and the concentration of different air pollutants and find a significant relationship between the concentration of particulate matter and sick-leaves, while the associations with SO_2 and NO_2 are more ambiguous.

Address: Anett C. Hansen, Statistics Norway, Research Department, P.O. Box 8131 Dep., N-0033 Oslo, Norway. E-mail: acn@ssb.no
Harald K. Selte, Norwegian Competition Authority, P.O. Box 8132 Dep., N-0033 Oslo, Norway. E-mail: Harald.K.Selte@nca.telemax.no

1. Introduction

1.1. Background

In recent years a growing number of studies have assessed the association between air pollution and human health. Significant exposure-response relations between ambient air pollution and different measures of morbidity and premature mortality have been documented in several studies (see below).

Health effects of air pollution may cause an indirect welfare loss to society in addition to pure economic costs. The pure economic costs consist of direct treatment costs and the efficiency losses due to an increased number of sick-leaves and a reduced labour force. The sick-leaves also implies fluctuations in the labour force which normally are considered costly. In addition there is an efficiency loss related to the fact that many will attend work although with less effort as they are affected. Several studies have calculated the costs associated with air pollution, see for instance Pearce and Crowards (1996) and Rosendahl (1996). However, these studies are based on transferring exposure-response functions from other countries. As the effect of air pollution on human health is complex, there is great uncertainty related to this method and studies based upon own data is useful.

We investigate the relationship between air pollution and the number of sick-leaves in Oslo. For this purpose we use air pollution data, provided by Oslo City Department of Environmental Health and Food Control (ODEH) and Norwegian Institute of Air Research (NILU), and temperature data, provided by Norwegian Meteorological Institute (DNMI), along with sick-leaves data from a major white-collar company in Oslo.

This paper is organised as follows: Below we give a short reference to the literature. In section 2 we present some general remarks about the association between air pollution and sick-leaves, based on the existing knowledge. We also discuss some other potential factors influencing the number of sick-leaves. In section 3 we describe the data and the statistical methods applied for estimating procedures while section 4 presents and discusses the results. Section 5 concludes.

1.2. Literature

The health effects which have been studied most extensively are related to short term response to pollution exposure. Studies have been performed on the connection between air pollution and premature mortality (Ostro, 1993), respiratory hospital admissions (Pope, 1991, Ransom & Pope, 1995), acute respiratory symptoms (Krupnick *et al*, 1990) and asthma attacks (Ostro *et al.,* 1991). See for instance ORNL/RFF (1994) and Clench-Aas & Krzyzanowski (1996) for a survey.

Few studies have focused on the impact of air pollution on sick leaves or other kinds of reduced labour productivity. If the adverse effects primarily affect the elderly, children or other people outside the labour force, these effects would be negligible. However, the literature indicates that even individuals in the labour force will be affected. Ostro (1983) discusses the association between total suspended particles (TSP) and *work loss days* (WLD) for those currently working, and *restricted activity days* (RAD) for the entire population. A *restricted activity day* is defined as a day where the individual alters her normal activity, without being absent from work. Ostro employs three different model specifications and finds significant correlation between TSP and both WLD and RAD in each specification, which indicates a robust association. Hausman *et al.* (1984) find a significant effect of TSP on WLDs. Both studies use data from the Health Interview Survey (HIS) of 1976 and calculate the concentration of TSP by using airport visibility data. Ostro (1987) and Ostro & Rothschild (1989) replicated the studies by using HIS-data from 1976 to 1981, while making use of fine particles ($PM_{2.5}$ - particulate matter with an aerodynamic diameter less than 2.5 μm) instead of TSP. Ostro (1987) finds a consistent association between fine particles and WLDs for the working labour force, and with RADs for the entire population. Ostro & Rothschild (1989) find that the concentration of fine particles is consistently associated with the number of restricted activity days (RAD) due to acute respiratory symptoms among workers aged 18 to 65. Zuidema and Nentjes (1997) have performed a study on

4

Dutch data from 29 administrative health districts. They use aggregated cross-sectional data on an annual basis for a three year period and their results show that the method applied highly influences the estimated coefficients. When OLS was applied they found a significant effect on WLS from sulphur dioxide (SO_2), black smoke (BS) and ammonia (NH_3), while applying a fixed effects method[1] gave a significant effect from sulphate aerosol (SO_4) and NH_3. Particles (TSP) was not significant in either of the specifications.

2. Air pollution, Exposure Assessment and Human Health Effects

Exposure
It is not straightforward to use the measurements of ambient air pollution as a measure of the individual exposure to air pollution. The monitored compound should instead be considered as an air quality indicator. The difficulties in assessing (individual) exposure on basis of outdoor monitoring sites/measurement of outdoor air pollution levels are pointed out in Clench-Aas & Krzyzanowski (1996)[2]:

«..Exposure to air pollution is a function of ambient pollution, of indoor concentrations both due to penetration of outdoor pollutants and indoor sources, and of the movements of the individuals from one micro-environment to another. Often only a limited number of outdoor monitoring sites provides estimates of exposure. Differences occur due to placement and number of outdoor measuring stations and reliability of measured data. Representativity of the stations for exposure of the population is not well described...»

Our hypothesis is that the employees are exposed to ambient air pollution on their way to and from work in the centre of Oslo, at work and at home. Members of the workforce are exposed to a high dose of air pollution for a short period of time every day along congested roads (in cars, or while waiting for public transport). The public transport stops are typically places with a relatively high level of air pollution (*hot spots*). In addition the employees will be exposed while at work. In a normal working environment, the indoor air will be filtered. At best this will be a so called EU 7 -filter, if any at all. This filter will still allow small particles to pass and hence, some of the outdoor pollution penetrates to the indoor environment. Thus there is significant work place exposure for small outdoor particles. For the same reason workers may possibly be exposed at home.

We assume that the short-term health effect of air pollution originates within seven days after exposure, and that the pollution level on a specific day does not affect the sick-leaves before the next day. For each pollutant we use the average pollution level over the last seven days with one day lag[3] as an explanatory factor in our model.

Expected effects from air pollution
The consequence for human health of ambient air pollution is complex and there is still much uncertainty as to how sick-leave is related to air pollution. We will briefly refer some of the conclusions from the literature concerning various pollutant's effect on human health. For a more comprehensive discussion of these effects we refer to the surveys mentioned above.

Particulate matter
Particulate matter is characterised by size (PM_{10}, $PM_{2.5}$, total suspended particles, etc.). Hence, the chemical composition can vary for a given concentration and fraction of particles. While the coarse fraction (particulate matter with diameter between 2.5 and 10 µm) generally originates from mechani-

---

[1] In short this method takes into account the fact that the relationship between WLD and pollution might differ across regions
[2] Pp.11.
[3] There are several missing observations in the air pollution data due to technical problems etc. We have allowed the average to be calculated when there are minimum 6 observations of the concentration the last seven days.

cal activities such as industrial processes, resuspension of road dust, etc., $PM_{2.5}$ (or the fine fraction) generally originates from incomplete combustion of fossil fuels. Even though the health effects of a given concentration of particles will depend on many factors, it is the size of the particles that is considered as important with respect to deposition in the respiratory tractus and therefore is a crucial determinant for health outcome (Trijonis, 1983). The deposition curve (of particulate matter in humans respiratory tractus) has a biphasic shape indicating that particulate matter smaller than $PM_{0.01}$ or greater than PM_3 have the highest percentage deposition in the tractus. In many cases, only PM_{10} concentrations are available.

Epidemiological studies of adverse health effects indicate a clear association with PM_{10}. For instance Gordian *et al.* (1996) study the association between PM_{10} and temperature with daily outpatient visits for respiratory illnesses including asthma, bronchitis (including non-specific cough) and upper respiratory illness (such as sore throat, earaches, sinusitis) among employees and dependants with comprehensive health insurance in Anchorage, Alaska. Their results show that a 10 $\mu g/m^3$ increase in PM_{10} resulted in a 3-6% increase in visits for asthma and a 1-3% increase in visits for upper respiratory diseases.

Nitrogen dioxide – NO₂
While health effects of NO_2 exposure have been demonstrated in many clinical studies, relatively few epidemiological studies have included NO_2 in their analysis. One explanation for this is that NO_2 has not been measured in town areas to the same extent as has been particulate matter and SO_2. In addition, epidemiological studies differ significantly from clinical in their exposure environment as they are related to NO_2 exposure in ambient air, where NO_2 is one of several pollution components whose exposures may be correlated (Ackermann-Liebrich *et al.*, 1996). This complex exposure situation makes it difficult to disclose the specific impact of NO_2 on human health. However, there exists some epidemiological evidence of the adverse health effects of NO_2 exposure, see e.g. Ackermann-Liebrich et al. (1996). Outdoor NO_2 stems primarily from combustion of fossil fuels.

Sulphur dioxide – SO₂
Sulphur dioxide is generated by the combustion of oil and coal and has until the last decade been considered one of the most severe pollutants in causing adverse health effects. It is possible that the observed effects of SO_2 in many instances have been reflecting the impact of particulate matter. Some recent results suggest though that SO_2 may have independent, additional health effect, but in these studies a distinction between the various pollutants' health impacts has not always been possible and hence the results are uncertain. During the last decade the level of SO_2 has been reduced considerably in Norway, implying that the concentration now is substantially below concentration levels in most of the areas where the studies have been conducted. (Se discussion below and Clench-Aas & Krzyzanowski, 1996 pp. 70-87). SO_2 is thus no longer considered a main contributor to pollution related health effects in Norway.

Ozone – O₃
Ozone is not included in the study because of lack of data. The explanation for this is that the study only covers the winter season when the concentration of ground-level ozone is low due to the climate in Oslo[4] (see SFT, 1992).

Threshold values
Several international studies of air pollution and health have been performed at low levels of exposure (see Clench-Aas & Krzyzanowski, 1996 for a documentation of each pollutant). Evidence suggests that also at low air pollution levels, even below the former WHO Air Quality Guidelines[5], adverse

---

[4] Ostro and Rothschild (1989) find no consistent association between ozone and RAD in their study. Yet they find some evidence for an association with *minor restricted activity days (MRAD)*. MRAD is defined as restricted activity days that do not result in work loss.

[5] In the Revised WHO Air Quality Guideline Report (1996) no guideline for either long-term or short-term average concentrations of particulate matter is recommended.

health effects occur. Especially for particulate matter the literature suggests that from the current data base no threshold level can be derived and that the association between particulate air pollution and adverse health effects is linear above the lowest average particle concentrations measured and up to the concentration levels normally encountered in Norway (Clench-Aas & Krzyzanowski, 1996). Considering this we have no threshold value for any of the pollutants in this study.

Temperature
Temperature is an important confounder in time series studies of the relationship between air pollution and sick-leaves in several ways: The morning temperature a given day can have an acute or psychological effect on the decision of staying home from work the same day. Another possibility is that it takes some time before a fall in temperature influences sick-leaves or that a fall in temperature must sustain for a period of time in order to have an impact. In addition it is possible that colder weather worsen air quality through inversion, and that there are synergistic effects on sick-leaves from temperature and humidity, and air pollution, e.g. for asthmatics.

Finally we note that other interaction effects also may be present in the relationship between sick-leaves and air-pollution, for instances synergistic effects in mixed pollution situations. As we do not have any prior knowledge about how such interaction effects eventually should be modelled, this will not be pursued further in this study.

3. Data and method

3.1. The data
ODEH and NILU measure daily the concentration of various air pollution components at several locations in Oslo. We use data from two stations located in the centre of the city (St. Olavs square and Nordahl Bruns street) that are protected against direct influence from traffic. The daily average concentration of different air pollutants is obtained daily during the winter season. We use data from the winter 90/91 season to the winter 95/96 season, referred to as period 1 to 6. The concentrations of NO_2 and SO_2 are measured in all periods. The series of particulate matter are incomplete. In period 1 to 4, black smoke (BS) was measured. The measure of particulate matter (PM_{10}) started in the fifth period. We have not converted the BS measures to PM_{10}. Both Dockery and Pope (1994) and Pearce and Crowards (1996) use (with reservations) a one to one relationship between BS and PM_{10} in their meta-studies. This equivalence is disputed. Black smoke is an optical gauge measuring the darkness of particles collected on a filter and determined by reflectance (Dockery and Pope, 1994). This measure is converted to $\mu g/m^3$ by using a European standard. PM_{10} is measured by weight. Thus, the equivalence depends on the darkness of the particulates. A major source of particulate matter in Oslo is the extensive use of studded tyres in the winter season. It seems that black smoke only covers parts of these particulates (see Pearce and Crowards, 1996). When estimating, we use black smoke in the first four periods, and PM_{10} in the 5th and 6th period. We introduce a dummy variable to distinguish the effect of BS and particulate matter on the intercept. Temperature, in C^o, is measured daily by DNMI. The morning temperature the same day (TEMP07) is measured at 7 a.m. and is meant to capture the acute effect, while average temperature the seven previous days with one day lag (AV-TEMP), is meant to capture the medium term effect. The weekly average is computed as an average of the daily temperatures at 7 a.m. and 1 and 7 p.m..

Table 1a, 1b and 2 show some statistics for the pollution variables. Note in particular the following: The mean concentration of BS was 26.7; for PM_{10} it was 24.5 $\mu g/m^3$. The means of NO_2 and SO_2 were 40.9 $\mu g/m^3$ and 10.8 $\mu g/m^3$, respectively. We see that the pollution of SO_2 was decreasing through the whole period of the study and is very low compared with the level in other epidemiological studies (as discussed above). All pollutants are positively correlated and we see the highest correlation between NO_2 and BS. The morning temperature is negatively correlated with all pollutants. The sick-leaves, measured in per cent of present employees, are positively correlated with all pollutants and negatively

correlated with the morning temperature. Note however that these correlation coefficients are calculated by using the daily means, and not the 7-day averages used in estimation.

Table 1a. Mean values of daily observations

Variable:	Mean	Std. dev	Max.	Min.
PM_{10} [1]	24.46	13.85	87.60	2.70
Black smoke [2]	26.72	21.26	155.00	1.00
NO_2	40.93	16.96	129.60	6.00
SO_2	10.79	7.81	53.00	0.00
Morning temp.	-1.03	5.18	14.00	-17.90
Sick-leave in %	4.03	0.81	8.21	2.45

[1] Based on data from season 5 and 6. [2] Based on data from season 1 to 4.
Sources: Statistics Norway, ODEH, NILU and DNMI.

Table 1b. Correlation between daily observations

	Correlation with:				
Variable:	BS	NO2	SO2	Morning temp.	Sick-leave
PM_{10} [1]	0.68*	0.64	0.42	-0.34	0.17
Black smoke [2]		0.79	0.54	-0.28	0.01
NO_2			0.48	-0.37	0.05
SO_2				-0.29	0.11
Morning temp.,C^o					-0,19

* Based on data from the experimental measurements, see the note in table 2
[1] Based on data from season 5 and 6. [2] Based on data from season 1 to 4
Sources: Statistics Norway, ODEH, NILU and DNMI.

Sick-leaves
The sickness benefits are regulated by law in Norway. Employees can be absent due to illness for a period up to three days without getting a doctor's statement, for a maximum of four times a year. For longer periods of illness a statement is needed. Full wage compensation is granted from the first day of illness, even with a self-statement.

The company providing us with sick-leaves data is located in the centre of Oslo, not distant from the air pollution monitoring sites. For each working day in the period from 1991 to January 1996 employees on sick-leave are counted. To sort out people absent for other reasons (e.g. employees on leave, vacation or taking care of sick members in their family) we also have daily counts on employees on duty. The number of observations in the periods that correspond to the air pollution data is 716 when the Christmas and Easter holidays are kept out. The average number of employees at work is about 1500[6], and the share of employees on sick-leave varies between 2.4 and 8.2 per cent with an average at 4.03 per cent.

Individual characteristics such as age, sex, education, income and smoking habits are presumably important factors in explaining reasons for sick-leaves. As this study is performed on aggregate data, we are restricted to estimate the air pollution coefficient excluding individual information. However, the estimated coefficient for the included variables will be consistent if the included variables and the omitted variables are independent. This result which refers to a linear regression model, is also valid in a multinomial logit model, see Lee (1982). If the number of workers is high, and the period is relatively short, we may assume that the group of workers is homogenous. The personal characteristics can then be treated as constant within the group, and hence independent of the included variables.

Table 2. Mean values of variables used in estimation**

Period:	Sick-	Present	Numb.of	PM_{10} in	Numb.	B. s. in	Numb.	NO_2 in	Numb.	SO_2 in	Numb.	Temp.

[6] Excluded from the tables.

	leave (%)	Work-ers	work-ingdays[1]	µg/m³ *	of obs.	µg/m³ *	of obs.	µg/m³	of obs.	µg/m³	of obs.	at 7 AM, C°
1. (Jan. 1. 91 - April 1. 91)	4.16	1338	57	28.23[2]	26[2]	35.60	88	49.90	84	17.49	82	-2.91
2. (Oct. 1. 91 - April 1. 92)	4.22	1468	124	33.26[2]	27[2]	28.33	159	39.57	177	12.13	159	0.78
3. (Oct. 1. 92 - April 1. 93)	3.79	1508	124	27.37[2]	26[2]	25.93	177	40.86	176	13.20	173	-0.65
4. (Oct. 1. 93 - April 1. 94)	4.42	1519	120	–	0	22.80	169	42.60	146	11.48	170	-2.85
5. (Oct. 1. 94 - May 1. 95)	4.03	1573	139	23.72	191	–	–	37.66	191	6.57	180	-0.09
6. (Oct. 1. 95 - Feb. 1. 96)	3.48	1676	82	25.65	120	–	–	39.87	111	6.37	122	-1.81
Jan. 1. 91 - Feb. 1. 96	4.03	1523	646	24.46[3]	311[3]	26.72	593	40.93	885	10.79	894	-1.03

* The measurement of black smoke was succeeded by PM10 in the fourth season

**Except daily mean tempareture

1 Except workingdays between Christmas and New Years eve and between Palm Sunday and Easter

2 Results from experimental measurements done in February in the respective years

3 Not counting the results from the experimental measurements

Sources: ODEH, NILU, DNMI and Statistics Norway.

3.2. Empirical specification

An employee may be either working or on sick-leave, since employees on vacation or leave etc. are not counted. Let $Y_{it} = 1$ if employee i is absent day t, and zero otherwise. N_t is the number of employees that should be on work day t, and m_t is the number of absent employees, that is

$$(1) \qquad m_t = \sum_{i=1}^{N_t} Y_{it}.$$

The proportion of employees on sick-leave is denoted $q_t = m_t/N_t$, and $p_{it} \equiv \Pr(Y_{it} = 1)$. With aggregate data several personal characteristics that may affect the probability of getting sick, such as age, sex, education, income and smoking status are omitted (cf. discussion above) and hence $p_{it} = p_t$. Moreover the probability is assumed to be a function of a vector z_t of independent variables and a parameter vector β which we want to estimate,

$$(2) \qquad p_t = g(z_t, \beta),$$

where $g(\cdot)$ is a (yet) undescribed distribution function. z_t contains lagged air pollution variables (particles, SO_2, NO_2), the temperature measures and a trend factor. The evidence of a health-effect from particles seems clear in the literature, so we expect to find a positive sign for this parameter. The evidence for SO_2 and NO_2 is more ambiguous, but a non-negative sign is assumed. The effect of temperature, both weekly the average and the morning temperature, is expected to be negative. There has been a decline in the sick-leaves ratio over the past years in Norway that may not be explained by a change in the ambient air pollution, so the parameter connected with the trend variable might be negative.

The model
In the logit model, we assume that the distribution function $g(\cdot)$ in (2) have a logistic distribution. The logit specification[7] is close to the normal distribution, except for the tails, and have pleasant properties in calculation. $g(\cdot)$ is then

$$(3) \qquad p_t = \Lambda(z_t\beta) \equiv \frac{1}{1 + \exp(-z_t\beta)},$$

where $\Lambda(\cdot)$ is the logistic cumulative density function (cdf). See appendix A for more details on the econometric specification of this model.

---

[7] Studies like Hausman *et al.* (1984), Ostro, (1987) and Ranson & Pope (1995), use a Poisson model for count data (see Cameron & Triverdi (1986) or Green (1993)). This model is appropriate when there are many zero-observations. However it is not suited for this study since there are none zero-observations in our data-set.

The marginal effect of an increase in an exogenous variable is given by Maddala (1983):

$$(4) \qquad \frac{\partial p_t}{\partial z_t} = \Lambda'(z_t\beta)\beta_i = \frac{e^{z_t\beta}}{(1+e^{z_t\beta})^2}\beta_i,$$

which of course is dependent of the vector z_t.

In the logit model presented above we assumed that the error term was independent over time. The assumption of time independence is restrictive when we consider that *i)* sick-leaves may often last for more that one day, *ii)* the effect of omitted variables like infectious diseases may be correlated over time. The Durbin-Watson statistic from the regression of the model above (where the possible serial dependence is ignored) is below any significance points for first-order autocorrelation (AR1) (in the range 0.4 to 0.5 dependent of the specification, see table 4). This supports a suspicion of an auto-correlated error-term. To account for this autocorrelation we add a stochastic term, η_t, in (3):

$$(5) \qquad \Lambda^{-1}(p_t) = \log\left(\frac{p_t}{1-p_t}\right) = z_t\beta + \eta_t,$$

where η_t follows a first-order autoregressive process, and the error term e_t is independent and homoscedastic:

$$(6) \qquad \eta_t = \rho\eta_{t-1} + e_t$$

Note that this relates to the "true" underlying relationship, not the observed. Hence, if we replace p_t by q_t we obtain

$$(7) \qquad \log\left(\frac{q_t}{1-q_t}\right) = z_t\beta + v_t,$$

where the error-term v_t is given by

$$(8) \qquad v_t = \eta_t + u_t = \rho\eta_{t-1} + e_t + \frac{\varepsilon_t}{p_t(1-p_t)}.$$

Here e_t and ε_t are independent and homoscedastic, but v_t is both autocorrelated and heteroscedastic, which complicates the regression. The variance of the last term, u_t, in v_t is the same as in the case without autocorrelation. In our sample the variation in both N_t and q_t is rather small, so the correction for heteroscedasticity gives small impact on the estimated parameters and their standard deviation.[8] Thus, the loss from estimating (7) with an autoregressive disturbance when the heteroscedasticity from the last part of (8) is ignored, is small.

We tried to approach the time dependence problem in a more rigorous way by using a two-state Markov chain model. However, the model requires information that the original data did not contain. The estimation of the model was performed on data from another firm with fewer employees and other shortcomings without yielding any new conclusions. See Hansen and Selte (1997) for more details.

---

[8] The model without autoregression was estimated when ignoring the heteroscedasticity with small impact on the estimates and the *t*-values, see table 4.

4. Estimation results and discussion

The variables used in estimation are defined and their mean values and standard deviation reported in table 3. The models are estimated with several choices of exogenous variables. The results are reported in table 4 with t-values in parentheses. First, the model is estimated with all of the exogenous variables which we have considered, including both the weekly mean and the current morning temperatures. The sample correlation between the two temperature measures is at 0.64, indicating some degree of multicolinearity. The exclusion of either one of the temperature measures is not creating any big disturbances in the estimates, and the distortions are less in the AR(1) model. Finally, we have excluded both NO_2 and SO_2 which are insignificant and of changing sign in the other specifications. In addition the results from a simple OLS on the logit model are reported as an indicator on the loss from ignoring the heteroscedastisity in the AR(1) estimation.

Table 3. Variables used in estimation[1]

Variable	Definition	Mean	Std. dev.
SL	Share of workers on sick-leave	0.0406	0.0084
PM_{10}	Concentration of PM_{10}[2,3].	24.22	7.49
BS	Concentation of black smoke[2,4]	27.12	14.60
DUM_{BS}	Dummy = 1 when BS is used, 0 when PM is used	0.65	
NO_2	Concentration of NO_2[2].	41.56	11.14
SO_2	Concentration of SO_2[2].	10.79	5.68
Av-Temp	Temperature[5] in C°.	0.20	4.08
Temp 07	Temperature at 7 AM in C°	-1.27	4.93
Trend	Trend variable, 1 in 91, 2 in 91/92, 3 in 92/93 etc.	3.68	

[1] The values may differ somewhat compared to the values from estimation, due to exclusion of missing observations.

[2] Average daily mean, last 7 days, 1 day lag, $\mu g/m^3$.

[3] Data from the two last periods are used in calculating the mean and standard deviation.

[4] Data from the four first periods are used in calculating the mean and standard deviation.

[5] Average daily mean, last 7 days, 1 day lag.

Sources: ODEH, NILU, DNMI and Statistics Norway.

Table 4. Estimates of the parametervector β

	Serial independence	Serial dependence	Serial independence	Serial dependence	Serial independence	Serial dependence	Serial independence	Serial dependence	Heteroscedasticity and autocorrelation ignored	
Constant	-3.15 (-38.59)	-3.08 (-21.64)	-3.15 (-38.48)	-3.08 (-21.68)	-3.20 (-41.28)	-3.15 (-22.19)	-3.33 (-42.93)	-3.30 (-23.08)	-3.16 (-38.97)	-3.28 (-43.78)
PM_{10}	0.0069 (3.53)	0.0056 (2.24)	0.0068 (3.48)	0.0056 (2.25)	0.0071 (3.66)	0.0065 (2.60)	0.0052 (2.71)	0.0032 (1.44)	0.0071 (3.56)	0.0057 (3.01)
BS	-0.0025 (-2.40)	-0.00093 (-0.67)	-0.0027 (-2.51)	-0.0009 (-0.65)	-0.0026 (-2.50)	-0.0012 (-0.85)	-0.00096 (-1.17)	-0.00098 (-1.05)	-0.0027 (-2.63)	-0.0012 (-1.53)
DUM_{BS}	0.18 (3.17)	0.14 (1.47)	0.18 (3.23)	0.13 (1.46)	0.19 (3.36)	0.16 (1.75)	0.24 (4.07)	0.20 (2.17)	0.16 (2.91)	0.20 (3.52)
NO_2	0.00018 (0.16)	-0.000074 (-0.04)	0.00033 (0.28)	-0.00011 (-0.07)	0.00051 (0.45)	0.00016 (-0.09)			0.0007 (0.61)	
SO_2	0.00064 (0.27)	-0.0035 (-1.03)	0.00023 (0.10)	-0.0034 (-1.01)	0.0017 (0.72)	-0.0013 (-0.38)			0.00052 (0.22)	
Av-temp	-0.0048 (-3.61)	-0.0096 (-2.49)	-0.0077 (-3.23)	-0.0095 (-2.50)			-0.0098 (-3.37)	-0.0077 (-2.03)	-0.0058 (-2.00)	-0.0098 (-3.56)
Temp 07	-0.0036 (-1.64)	0.00035 (0.23)			-0.0058 (-3.24)	-0.00038 (-0.25)	-0.0024 (-1.08)	0.00052 (0.37)	-0.0021 (-0.98)	-0.0014 (-0.66)
Trend	-0.042 (-3.87)	-0.05 (-2.17)	-0.042 (-3.91)	-0.05 (-2.17)	-0.038 (-3.62)	-0.047 (-2.02)	-0.0013 (-0.13)	-0.053 (-0.21)	-0.049 (-4.68)	-0.018 (-1.76)
r		0.74 (24.18)		0.74 (24.16)		0.75 (24.40)		0.79 (30.16)		
n (days)	476	476	476	476	476	476	560	560	476	560
R^2-adj	0.103		0.10		0.10		0.11		0.12	0.10
DW-statistic	0.52	2.25	0.52	2.25	0.53	2.25	0.39	2.39	0.54	0.44

Dependent variable: Sick-leaves in share of the total number of possible man-days.
Heteroscedasticity is ignored in the autoregressive models.

The estimates from the logit estimations seem reasonable compared to the hypothesis in Section 2. Most of the interesting results are stable with respect to changes in model specification and the choice of independent variables. The R-squared may seem low (about 0.1), but this is not unexpected since we use daily data without individual characteristics. The Durbin-Watson statistic is low in the regressions without autoregressive error term (about 0.4-0.5). Disregarding the heteroscedastisity is not affecting the estimates particularly; most estimates are somewhat lower in absolute terms, and most *t*-values are slightly smaller in value, but the conclusions would not be altered to any great extent if these estimates were used. Thus, we expect that neglecting the heteroscedastisity in the AR(1) model only will have a small impact on the estimates and the *t*-values. In the AR(1) regression the estimates of the autocorrelation parameter ρ are high (0.75) and significant at all levels, and the DW-statistic of the transformed residuals is above 2, which both justify the assumption of time dependence.

PM_{10} shows a positive, significant and fairly stable relation with sick-leaves in all specifications, except in one of the AR(1) regressions. When black smoke is used as an indicator of the concentration of ambient particulate matter we find an unexpected negative effect, which also is insignificant in the AR(1) models and when NO_2 and SO_2 are excluded. The estimated parameter value for PM_{10} is about 0.0055 to 0.0070. In the regression of the AR(1) model where NO_2 and SO_2 are excluded the coefficient for PM_{10} is nearly halved (0.0039), with a *t*-value below 2. The effect on the predicted probability of getting sick of marginal changes in the exogenous variables is given by (analogous to eq. (4)):

$$(9) \qquad \frac{\partial \hat{p}_t}{\partial z_{it}} = \Lambda'(z_t\hat{\beta})\hat{\beta}_i = \frac{e^{z_t\hat{\beta}}}{(1+e^{z_t\hat{\beta}})^2}\hat{\beta}_i,$$

where $\hat{\beta}$ is the estimated parameter vector. The effect depends on the values of the exogenous variables, but as the constant term is dominating in $z_t\hat{\beta}$, the effects of changes in the exogenous are rather small. The marginal effects are calculated for PM_{10} and weekly mean and morning temperature at average levels of the independent variables; see table 5.

Table 5. Marginal effect on sick-leaves of marginal changes in the independent variables Calculated at mean level of all independent variables*

	All variables		Temp 07 excluded		Av-temp excluded		NO_2, SO_2 excluded	
	Serial independence	Serial dependence	Serial independence	Serial dependence	Serial independence	Serial dependence	Serial independence	Serial dependence
PM_{10}	0.00026	0.0002	0.00025	0.0002	0.00027	0.00023	0.00019	*0.00011* [n.s.]
Av-temp., C^o	-0.00019	-0.00037	-0.0003	-0.00036			-0.00039	-0.00029
Temp07, C^o	*-0.00014* [n.s.]	n.s.			-0.00023	n.s.	n.s.	n.s.

n.s.: not significant*For PM10, mean level in 94/95, 95/96 only.

We see from table 5 that an increase in the PM_{10}-concentration by one unit ($\mu g/m^3$) on average leads to a predicted increase in the sick-leaves in the range 0.019 - 0.027 percentage points[9] (because $\frac{\partial \hat{m}_t}{\partial z_{it}}\Big/N_t = \frac{\partial p_t}{\partial z_{it}}N_t\Big/N_t = \frac{\partial p_t}{\partial z_{it}}$) with an unweighted average of 0.0023. The effect on the sick-leaves ratio is an increase by 0.58 per cent when the average share of employees on sick-leaves is 4.0 per cent. As an example; in a firm with 1000 employees, the impact of an increase in the average PM_{10}-concentration during the season by 10 units is about 2.3 more sick-leaves in the firm every day in the period. That is 276 extra days of sick-leave during a winter season of 120 working days. Moreover, halving the average PM_{10} concentration in Oslo from 24.5 $\mu g/m^3$ to 12.3 $\mu g/m^3$ could reduce the sick-leaves ratio by 7 per cent.

The effects of NO_2 and SO_2 are both insignificant in all specifications. The epidemiological evidence of effects of NO_2 on morbidity is more uncertain than the evidence of particulate matter, as discussed above. Clench-Aas and Krzyzanowski (1996) have evaluated several studies that show adverse health effects of SO_2. They find no evidence of threshold values for SO_2, but all studies considered in their report are performed at levels of SO_2 concentration which by far exceeds the level in this study (average level at 10.8 $\mu g/m^3$). If we assume adverse health effects of SO_2 at high concentration levels; the reason for not finding any effect in our study at this low level of SO_2-concentration, might be non-linearity in the exposure-response relation of SO_2. That is that the exposure level either is below a threshold value for employed persons, or that the effect of marginal changes in the concentration is smaller (not measurable) at this low level.

Since NO_2 and SO_2 both are correlated with PM_{10}, omitting these variables affects the variance for the estimated PM_{10}-coefficient. That is, in model 4 we cannot preclude that the insignificant coefficient for PM_{10} results from omitting NO_2 and SO_2.

The two measures of temperature are used to distinguish between a possible acute effect (morning temperature) and a short- to medium term effect (weekly mean). The morning temperature is insignificant except when the weekly mean temperature is excluded and the autocorrelation is disregarded. Thus we find no support for the hypothesis of an acute effect of the morning temperature on the sick-leaves. The weekly mean temperature however shows a significant negative effect on the sick-leaves

---

[9] The non-significant estimate is disregarded.

in all estimations where it is included. This may be an indicator that the lower the temperature, the higher probability of getting ill (for example catching a cold) and being absent from work.

The trend variable is negative and significant except when NO_2 and SO_2 were excluded. A negative sign was expected, since the number of sick-leaves was decreasing in Norway from 1992 to 1995, (see NHO (1996) for short term and RTV (1996) for longer term sick-leaves).

5. Concluding remarks

The scope of this paper has been to study the relationship between air pollution and the number of sick-leaves as several studies have documented adverse health effects from air pollution. There are few internationally published studies that examine health effects of air pollution on the working population separately, but both Ostro and Rothschild (1989) and Gordian *et. al.* (1996) find evidence for adverse health effects of ambient particulate matter on this part of the population. Our study on Norwegian data confirms that air pollution has adverse productivity effects on the labour force.

We find a significant association between small particulate matter (PM_{10}) and the number of sick-leaves. An increase in the average level of PM_{10} by 1 $\mu g/m^3$ leads to an increase in the number of sick-leaves by about 0.6 per cent. When using black smoke as a particle indicator, no significant effect was identified. No significant effect was found neither for nitrogen dioxide (NO_2) nor sulphur dioxide (SO_2). However, the concentration of SO_2 in Oslo is very low compared to other cities where such adverse health effects of SO_2 are reported. Neither Ostro & Rothschild (1989) nor Gordian *et. al.* (1996) make use of NO_2 or SO_2 measures. In addition to general welfare losses (decline in the individual utility of being) and costs through increase in public health services, our results indicate that air pollution also brings about costs to trade and industry, through higher sick-leaves.

References

Ackermann-Liebrich, U., G. Pershagen and R. Rapp (1996): Epidemiological Evidence of Health Effects of Nitrogen Dioxide, Paper presented at MRC/IEH Institute for Environment and Health (1996) Workshop on Effects of Ozone and Nitrogen Oxides in an Integrated Assessment of Air Pollution, Eastbourne, UK 10-12. June 1996.

Amemiya, T. (1985): *Advanced Econometrics*, Basil Blackwell, Oxford.

Cameron, A.C. and P.K.Triverdi (1986): Econometric Models based on Count Data: Comparisons and Applications of some Estimators and Tests, *Journal of Applied Econometric* **1**, 29-53.

Clench-Aas, J. and M. Krzyzanowski (eds.) (1996): *Quantification Of Health Effects Related to SO₂, NO₂, O₃ and Particulate Matter Exposure*, Report from the Nordic Expert Meeting Oslo, NILU/WHO, 15-17 Oct. 1995, Kjeller: Norwegian Institute for Air Research and Bilthoven: WHO Regional Office for Europe.

Dockery, D.W. and C.A. Pope III (1994): Acute respiratory effects of particulate air pollution, *Annual review of Public Health* **15**, 107-132.

Gordian, M.E., H. Özkaynak, J. Xue, S.S. Morris and J.D. Spengler (1996): Particulate Air Pollution and Respiratory Disease in Anchorage, Alaska, *Environmental Health perspectives* **104**,3, 290-297.

Greene, W.H. (1993): *Econometric analysis*, second edition, Macmillian, New York.

Hansen, A.C. and H.K. Selte (1997): Applying a two-state Markov chain model for the relationship between air pollution and sick-leaves. To be published in the series Documents, Statistics Norway.

Hausmann, J.A., B. Ostro and D.A. Wise (1984): Air Pollution and Lost Work, Working Paper 1263, National Bureau of Economic Research (NBER), Cambridge.

Krupnick,A.J, W. Harrington and B. Ostro (1990): Ambient Ozone and Acute Health Effects: Evidence from Daily, Journal of Environmental Economics and Management **18**, 1-18.

Lee, L.-F. (1982): Specification error in multinomial logit models, Analysis of the Omitted Variable Bias, *Journal of Econometrics* **20**, 197-209.

Maddala, G.S. (1983): *Limited-dependent and qualitative variables in econometrics*, Cambridge University Press, Cambridge.

NHO (1996): *Lønns og fraværs statistikk, 2. kvartal 1996* (Wage and absence statistics, 2nd quarter 1996), Næringslivets Hovedorganisasjon, Oslo.

Oak Ridge National Laboratory and Resources for the Future (ORNL/RFF) (1994): *External Costs and Benefits of Fuel Cycles: A Study by the U.S. Department of Energy and the Commission of the European Communities*, Oak Ridge, TN.

Ostro, B. (1983): The Effects of Air Pollution on Work Loss and Morbidity, *Journal of Environmental Economics and Management* **10**, 371-382.

Ostro, B. (1987): Air Pollution and Morbidity Revisited: A Specification Test, *Journal of Environmental Economics and Management* **14**, 87-98.

Ostro, B. (1993): The association of Air Pollution and Mortality: Examining the Case for Inference, *Archives of Environmental Health* **48**, 336-342.

Ostro, B. and S. Rothschild (1989): Air Pollution and Acute Respiratory Morbidity: An Observational Study of Multiple Pollutants, *Environmental Research* **50,** 238-247.

Ostro, B., M.J. Lipsett, M.B. Wiener and J.C. Selner (1991): Asthmatic Responses to Airborne Acid Aerosols, *American Journal of Public Health* **81**, 694-702.

Pearce, D. and T. Crowards (1996): Particulate matter and human health in the United Kingdom, *Energy Policy* **24**, 609-619.

Pope, C.A. III (1991): Respiratory Hospital Admissions Associated with PM_{10} Pollution in Utah, Salt Lake and Cache Valleys, *Archives of Environmental Health* **47**, 211- 217.

Ransom, M.R. and C.A. Pope III (1995): External Health Costs of a Steel Mill, *Contemporary Economic Policy* **13**, 86-97.

Rosendahl, K. E. (1996): *Helseeffekter av luftforurensning og virkninger på økonomisk aktivitet* (Health effects of air pollution and effects on economic activity*)*, Reports 96/8, Statistics Norway, Oslo.

RTV (1996), *Trygdestatistisk årbok 1996* (Statistical yearbook of the National Health Insurance 1996), Rikstrygdeverket, Oslo 1996.

SFT (1992): Virkninger av luftforurensninger på helse og miljø- anbefalte luftkvalitetskriterier, (Effects of air pollution on health and environment - air quality guidelines) Rapport 92:16.

Trijonis, J. (1983): Development and Application of Methods for Estimating Inhalable and Fine Particle Concentrations from Routine Hi-Vol Data, *Atmospheric Environment* **17**, 999-1008.

Zuidema, T. and A. Nentjes (1997): Health Damage of Air Pollution: An Estimate of a Dose-Response Relationship for the Netherlands, *Environmental and Resource Economics* **9**, 291-308.

Appendix A

The logistic cdf, $\Lambda(\cdot)$, is overall increasing so we may write:

$$(A1) \qquad \Lambda^{-1}(p_t) = \log\left(\frac{p_t}{1 - p_t}\right) = z_t\beta .$$

We know that q_t is an unbiased and consistent estimator of p_t, that is $q_t = p_t + \varepsilon_t$, or $\Lambda^{-1}(q_t) = \Lambda^{-1}(p_t + \varepsilon_t)$ where ε_t is a random term with $E(\varepsilon_t) = 0$. First order Taylor expansion yields $\Lambda^{-1}(p_t + \varepsilon_t) \approx \Lambda^{-1}(p_t) + \varepsilon_t/(p_t(1-p_t))$. So from eq. (A1) we get:

$$(A2) \qquad \Lambda^{-1}(q_t) = \log\left(\frac{q_t}{1 - q_t}\right) = z_t\beta + u_t,$$

where $u_t = \varepsilon_t/(p_t(1-p_t))$, $E(u_t) \approx 0$ and the variance of u_t is given by

$$(A3) \qquad \text{var}(u_t) \approx \frac{1}{N_t p_t (1 - p_t)},$$

so u_t is heteroscedastic. The variance can be estimated by substituting p_t by q_t in (A3) (Maddala 1983, Amemiya 1985) or by $\hat{p}_t = \Lambda(z_t\tilde{\beta})$, where $\tilde{\beta}$ is the OLS-estimate of β from eq (A2) (Greene 1993). β can be estimated by Berkson's minimum chi-squared method (See Maddala, 1983), that is to use weighted least squares on (A2) with $\Lambda^{-1}(q_t)$ as dependent variable, using the estimated variance of u_t as weights. This procedure produces consistent estimates given the assumptions.

Recent publications in the series Discussion Papers

153 A. Langørgen (1995): On the Simultaneous Determination of Current Expenditure, Real Capital, Fee Income, and Public Debt in Norwegian Local Government

154 A. Katz and T. Bye(1995): Returns to Publicly Owned Transport Infrastructure Investment. A Cost Function/ Cost Share Approach for Norway, 1971-1991

155 K. O. Aarbu (1995): Some Issues About the Norwegian Capital Income Imputation Model

156 P. Boug, K. A. Mork and T. Tjemsland (1995): Financial Deregulation and Consumer Behavior: the Norwegian Experience

157 B.E. Naug and R. Nymoen (1995): Import Price Formation and Pricing to Market: A Test on Norwegian Data

158 R. Aaberge (1995): Choosing Measures of Inequality for Empirical Applications

159 T.J. Klette and S.E. Førre (1995): Innovation and Job Creation in a Small Open Economy: Evidence from Norwegian Manufacturing Plants 1982-92

160 S. Holden, D. Kolsrud and B. Vikøren (1995): Noisy Signals in Target Zone Regimes: Theory and Monte Carlo Experiments

161 T. Hægeland (1996): Monopolistic Competition, Resource Allocation and the Effects of Industrial Policy

162 S. Grepperud (1996): Poverty, Land Degradation and Climatic Uncertainty

163 S. Grepperud (1996): Soil Conservation as an Investment in Land

164 K.A. Brekke, V. Iversen and J. Aune (1996): Soil Wealth in Tanzania

165 J.K. Dagsvik, D.G. Wetterwald and R. Aaberge (1996): Potential Demand for Alternative Fuel Vehicles

166 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part I: Theory

167 J.K. Dagsvik (1996): Consumer Demand with Unobservable Product Attributes. Part II: Inference

168 R. Aaberge, A. Björklund, M. Jäntti, M. Palme, P. J. Pedersen, N. Smith and T. Wennemo (1996): Income Inequality and Income Mobility in the Scandinavian Countries Compared to the United States

169 K. Nyborg (1996): Some Norwegian Politicians' Use of Cost-Benefit Analysis

170 E. Berg, S. Kverndokk and K. E. Rosendahl (1996): Market Power, International CO_2 Taxation and Petroleum Wealth

171 R. Aaberge, U. Colombino and S. Strøm (1996): Welfare Effects of Proportional Taxation: Empirical Evidence from Italy, Norway and Sweden

172 J.K. Dagsvik (1996): Dynamic Choice, Multistate Duration Models and Stochastic Structure

173 J.K. Dagsvik (1996): Aggregation in Matching Markets

174 H.C. Bjørnland (1996): The Dynamic Effects of Aggregate Demand, Supply and Oil Price Shocks

175 A. Bruvoll and K. Ibenholt (1996): Future Waste Generation. Forecasts Based on a Macroeconomic Model

176 T. Fæhn and L. A. Grünfeld (1996) Recent Leaps Towards Free Trade. The Impact on Norwegian Industry and Trade Patterns

177 R. Barrell and K. A. Magnussen (1996): Counterfactual Analyses of Oil price Shocks using a World Model

178 E. Bowitz and S. I. Hove (1996): Business cycles and fiscal policy: Norway 1973-93

179 H.C. Bjørnland (1996): Sources of Business Cycles in Energy Producing Economies: The case of Norway and United Kingdom

180 K. Nyborg (1996): The Political Man and Contingent Valuation: Motives Do Count

181 E. Berg, S. Kverndokk and K.E. Rosendahl (1996): Gains from Cartelisation in the Oil Market

182 R. Aaberge and I. Aslaksen (1996): Decomposition of the Gini Coefficient by Income Components: Various Types of Applications and Interpretations

183 B. Bye (1996): Taxation, Unemployment and Growth: Dynamic Welfare Effects of "Green" Policies

184 T.J. Klette and F. Johansen (1996): Accumulation of R&D Capital and Dynamic Firm Performance: A Not-so-fixed Effect Model

185 B. Bye (1996): Environmental Tax Reform and Producer Foresight: An Intertemporal Computable General Equilibrium Analysis

186 S. Grepperud (1997): Soil Depletion Choices under Production and Price Uncertainty

187 N.M. Stølen and T. Åvitsland (1997): Has Growth in Supply of Educated Persons Been Important for the Composition of Employment?

188 T.J. Klette and Z. Griliches (1997): Empirical Patterns of Firm Growth and R&D Investment: a Quality Ladder Model Interpretation

189 J. Aune, S. Glomsrød, V. Iversen and H. Wiig (1997): Structural Adjustment and Soil Degradation in Tanzania. A CGE-model Approach with Endogenous Soil Productivity

190 E. Biørn and T.J. Klette (1997): Panel Data with Errors-in-Variables: A Note on Essential and Redundant Orthogonality Conditions in GMM-estimation

191 L. Belsby and B.K. Wold (1997): Primary Schooling in Zambia Squeezed at Community and Household Level

192 E. Bowitz and Å. Cappelen (1997): Incomes Policies and the Norwegian Economy 1973-93

193 S. Glomsrød, M.D Monge A. and H. Vennemo (1997): Structural Adjustment and Deforestation in Nicaragua

194 F. Johansen and T.J. Klette (1997): Wage and Employment Effects of Payroll Taxes and Investment Subsidies

195 T. Fæhn (1997): Non-Tariff Barriers - the Achilles' Heel of Trade Policy Analysis

196 R. Aaberge and A. Langørgen (1997): Fiscal and Spending Behavior of Local Governments: An Empirical Analysis Based on Norwegian Data

197 A.C. Hansen and H.K. Selte (1997): Air Pollution and Sick-leaves - is there a Connection? A Case Study using Air Pollution Data from Oslo