

**Tjalling C. Koopmans Research Institute**

*Tjalling C. Koopmans*



**Universiteit Utrecht**

**Utrecht School  
of Economics**

**Tjalling C. Koopmans Research Institute  
Utrecht School of Economics  
Utrecht University**

Vredenburg 138  
3511 BG Utrecht  
The Netherlands  
telephone +31 30 253 9800  
fax +31 30 253 7373  
website [www.koopmansinstitute.uu.nl](http://www.koopmansinstitute.uu.nl)

The Tjalling C. Koopmans Institute is the research institute and research school of Utrecht School of Economics. It was founded in 2003, and named after Professor Tjalling C. Koopmans, Dutch-born Nobel Prize laureate in economics of 1975.

In the discussion papers series the Koopmans Institute publishes results of ongoing research for early dissemination of research results, and to enhance discussion with colleagues.

Please send any comments and suggestions on the Koopmans institute, or this series to [M.Damhuis@econ.uu.nl](mailto:M.Damhuis@econ.uu.nl)

ontwerp voorblad: WRIK Utrecht

**How to reach the authors**

*Please direct all correspondence to the first author.*

Wolter Hassink  
Utrecht University  
Utrecht School of Economics  
Vredenburg 138  
3511 BG Utrecht  
The Netherlands.  
Email: [w.hassink@econ.uu.nl](mailto:w.hassink@econ.uu.nl)

Pierre Koning  
Utrecht University  
Utrecht School of Economics  
Vredenburg 138  
3511 BG Utrecht  
The Netherlands.  
E-mail: [p.koning@econ.uu.nl](mailto:p.koning@econ.uu.nl)

## **Do Financial Bonuses to Employees Reduce their Absenteeism? Outcome of a Lottery**

Wolter Hassink  
Pierre Koning

Utrecht School of Economics  
Utrecht University

June 2005

### **Abstract**

This paper investigates the effectiveness of a monthly lottery in reducing sick leave among workers in a manufacturing firm. Conditions of participation are not having reported sick in the previous three months and not having won the lottery earlier. It turns out that the lottery results in a decrease in the rate of sick leave of 1.6 percentage point. From the perspective of the firm, the lottery is found to be highly beneficial – that is, the benefits associated with the decrease in the sick leave rate exceed the costs of the lottery. Workers seem to be primarily driven by the first upcoming lottery. After winning the lottery, winners resume their previous (rate of) absence.

**Keywords:** absenteeism; sick leave; incentives; lottery

**JEL classification:** J22; J32; M52

### **Acknowledgements**

We acknowledge Pauline Altena for her indispensable help in acquiring the data as well as the anonymous firm for kindly giving us access to its personnel records. We have benefited from comments of Iwan Barankey, Tim Barmby, Giovanni Russo, Kathryn Shaw, and seminar participants at the 2<sup>nd</sup> SOLE/EALE World Meeting 2005.

## **1. Introduction**

At present, there is strong empirical evidence that worker absenteeism behavior responds to negative financial incentives. First, a well-known empirical finding is that absenteeism is inversely related to the unemployment rate. Since absence-prone employees are more likely to be laid off, the threat of being laid off has a disciplining effect and employees are less likely to shirk when unemployment is high (see e.g. Leigh (1985) and Kaivanto (1997)). Second, there is empirical evidence that employees have reduced rates of absenteeism when they experience a wage decrease upon absence (e.g. Barmby et al. (1995); Johansson and Palme (1996); Henrekson and Persson (2004)).

One common feature that drives the outcomes of these empirical studies is that absenteeism is associated with negative incentives for the employee. Yet, in many employer-employee relationships in various European countries, negative incentives for absenteeism are not an easy option due to income and job protection for the absentees. For instance, although Dutch employers are obliged to continue to pay only 70 percent of gross earnings during sickness, most of the collective bargaining agreements between employers and employees stipulate that employers supplement the benefits to net earning levels (De Jong and Lindeboom (2003)). Income protection for absenteeism is not a typically Dutch phenomenon. Also in surrounding countries, like Germany and Denmark, most of the employees do not experience an immediate wage drop when being on sick leave. Sickness benefits are usually topped up by extra-legal-arrangements, which have been laid down in (collective) labor contracts (Einerhand et al. (1995); Barmby et al. (2002)).

When firms cannot apply negative incentives for reasons of income protection, they may use bonuses instead to attain a lower rate of sickness absence. So far, few empirical studies have investigated the impact of incentive pay on absenteeism (Wilson and Peel (1991); Brown et al. (1999); Engellandt and Riphahn (2004)). All of these studies find incentive pay to reduce absenteeism, but reduced absenteeism is only one of the mechanisms that increase the productivity of workers – or, stated differently, incentive pay is not directly linked to absenteeism behavior.

By contrast, the attendance bonus system we will investigate in this paper is

related directly to individual sick leave behavior. We analyze a unique monthly lottery set up in two plants (with in total about 400 employees) of a large Dutch manufacturer. Participation in the lottery is related to the incidence of sick leave in the past months, with the intention to stimulate workers to reduce their absenteeism. Only those employees who had not been on sick leave for the past three months were eligible for the lottery. Each month, seven randomly selected winning workers received a gift coupon of 75 euros. Winners were excluded from subsequent lotteries. Hence, the setup of the lottery makes it possible to use information on sick leave in three stages of the lottery: the period prior to the lottery; during the lottery when a worker has not been selected a winner; and in the period after a worker had won the lottery. The aim of this paper is to analyze empirically whether the financial incentives associated with the lottery have resulted in a lower rate of sick leave.<sup>1</sup>

The outline of this paper is as follows. Section 2 discusses the literature on absenteeism and incentives. Section 3 describes the lottery in great detail. Section 4 presents the empirical model, and in Section 5 the estimation results are discussed. Section 6 presents the conclusion.

## **2. Literature on absenteeism and incentives**

Economic studies of sick leave and workplace absence are based on the assumption that employers can only imperfectly monitor their employees' activities. When employees regard work as a disutility, they may have a higher incidence of absenteeism than is indicated by their health status.<sup>2</sup> The rate of absence is affected by the costs incurred by both the employer and the employee. When unanticipated sick leaves are costly to employers due to foregone production, employees with lower rates of absenteeism may be compensated through higher (hedonic) wages (e.g. Allen (1981)). For employees, higher costs attached to sick

---

<sup>1</sup> A complete evaluation of the incentive system should take account of the externality effects. The incentive system may stimulate workers not to report sick, with the unintended consequence that they infect their co-workers on the work floor; i.e. workers do not internalize the external effects of their sick leave decision. Consequently, the incentive system may lead to an increase in absenteeism. We leave this issue unanswered.

<sup>2</sup> We do not pay attention to the impact of the flexibility of the contract with respect to working time (see e.g. Allen (1981); Dunn and Youngblood (1986); Kenyon and Dawkins (1989)).

leaves may stimulate them to reduce their absenteeism. We distinguish three types of incentives that have been reported in the literature.

The first class of studies addresses the employees' responses to an immediate wage decrease upon absence. Winkler (1980) shows that income protection against sick leave is associated with a higher absence among teachers. Doherty (1979) demonstrated that in the U.K. higher sickness benefits resulted in moral hazard with higher sick leaves. Studies into the social insurance system in Sweden have a similar outcome (Johansson and Palme (1996, 2002); Henrekson and Persson (2004)). Next to incentives due to nationwide or sectoral insurance systems, employees may be confronted with firm-specific incentive schemes. Barmby et al. (1991, 1995) investigated the consequences of such a scheme for a large industrial firm. In the scheme, workers get points in case of "unacceptable" sickness that has not been certified by a general practitioner or hospital. After two years, workers may receive a bonus or a sanction, depending on the number of points they have gathered. Barmby et al. demonstrated that this scheme led to lower absenteeism.

The second class of studies focuses on future consequences of absenteeism. With frequent absenteeism it may become more unlikely that employees get promoted or that they receive merit pay. They may even get fired when the employer discovers them shirking. Consequently, and following standard efficiency wage theory arguments, as monitoring costs are higher in large firms, the rate of absence is positively related to firm size (Winkler (1980); Winkelmann (1999)). Similarly, since the costs of shirking are higher in slack labor markets due to increased layoff probability, the rate of absence is negatively related to the unemployment rate (e.g. Leigh (1985); Kaivanto (1997)). Shirking is particularly costly for temporary workers, as firing costs are low here (see e.g. Arai and Thoursie (2005)). Ichino and Riphahn (2004) demonstrated that white-collar workers who were hired by a large Italian bank had a lower absence rate during their probation period, when they were unprotected against dismissal.

The third class of studies investigates whether positive incentives through rewards and bonuses result in a better performance of absenteeism. Probably the most prominent examples of such (positive) incentive schemes at the group level

are employee profit sharing and shared ownership. Although these schemes do not primarily aim at reducing sick leave, they may lead to improved work conditions, higher morale, better motivation and more job satisfaction – thus decreasing the sick leave rate. Wilson and Peel (1991) (for U.K. firms) and Brown et al. (1999) (for French firms) showed that employee sharing leads to lower absence rates. Engellandt and Riphahn (2004) investigated the effect of performance-related pay for employees of a Swiss organizational unit of an international company. They found that higher pay leads to lower absence rates.

### **3. The lottery**

The incentive scheme for sick leave that we will investigate is aimed at workers in two plants with capital-intensive technologies. Machines are operated 24 hours a day, except for weekends and public holidays. Workers are employed in shifts, in order to keep the machines running. Consequently, absenteeism is expensive to the firm, since it needs to recruit additional workers at considerable costs in case of severe shortages of labor. On July 1<sup>st</sup> 2001, both plants were acquired by a large Dutch manufacturing firm. In the first months, the manufacturer reorganized the departments and changed some of the job titles. The number of workers declined by about 10 percent from 419 to 376 workers, from July 1<sup>st</sup> 2001 to July 31<sup>st</sup> 2003.

Shortly after the acquisition, the daily rate of sick leave was high, ranging between 10 and 15 percent. In April 2002, it induced the firm to announce to its workers a monthly lottery, starting at June 2002. The lottery was based on the following scheme. At the beginning of each month, the firm selects the workers who have not been on sick leave in the previous three months.<sup>3</sup> From this group of workers, seven winners are selected randomly, for which the firm uses two sampling strata: Three winners in the smaller plant (with about 40% of the total number of workers) and four in the other plant. Each lottery winner receives a coupon gift with a value of 75 euros. Furthermore, the names of the winners are made publicly known through the entire company. Winners are excluded from

---

<sup>3</sup> In June 2002 (July 2002), the restriction pertained to workers who had not been on sick leave in the past (two) month(s).

future lotteries, even if they have not been on sick leave during a period of three months. Hence, lottery winners lose an incentive to reduce their sick leaves.

Our data were obtained from the individual personnel records of all workers for the period July 1<sup>st</sup> 2001 – July 31<sup>st</sup> 2003. It results in an unbalanced panel dataset of 481 workers; over the period of investigation, the firm hired 62 workers and 105 workers left the firm. Furthermore, we observed the absence records of these workers at a daily basis over the same period. As the lotteries were organized on a monthly basis, we have transformed the daily information into a dummy variable that measures the monthly incidence of sick leave. It equals one if the worker was on sick leave at any of the working days during the month, and zero if otherwise. In addition, we calculated the individual duration of sick leave by constructing a variable that registers the fraction of working days for which the worker was on sick leave during the month. Thus, the sick leave data were aggregated over time from a daily basis to a monthly basis, which resulted in a data set of 9,637 monthly observations.<sup>4</sup>

From this broad data set we selected two samples. First, for all workers we excluded the monthly observations in which a worker was on sick leave on all working days of the previous month. As we will motivate more extensively in Section 4, workers who were temporarily in bad health are not able to respond to the lottery anyway. We claim that these observations are no part of the population of workers for whom the lottery was set up. Thus, we exclude these monthly observations, yielding a selected sample of 9,088 (monthly) observations. Second, we used a selected sample of employees consisting of all of the winners. In all, the number of monthly observations of the winners added up to 2,183, with 2,116 with no sick leave on all working days of the previous month. The empirical analyses in Section 5 are based on these two data sets.

*<Table 1>*

We discuss the observable characteristics of the workers for both subsamples. In Table 1 we distinguish between the (gross) population and the selection

---

<sup>4</sup> For all of these monthly observations, the worker was employed with the firm at all of the working days of the calendar month. Consequently, workers who were hired during a month (or left the firm during a month) are included in the sample from the first calendar month after hiring (up to and including month before separation).



of workers who were not on sick leave on all working days of the previous month. Furthermore, the table presents two sick leave measures. First, the monthly incidence of sick leave is defined as the occurrence of reporting sick in a particular month, irrespective of the number of sick leave days. This variable is (primarily) used in the estimation of the model, as the lottery incentive is based on the incidence of reporting sick in a particular month as well. Second, the monthly duration or the rate of sick leave is defined as the number of days a worker reported sick in a particular month, and is expressed as a percentage of the number of working days. Obviously, and in contrast to the monthly incidence of sick leave, this variable is informative on the lottery incentive effect in terms of reduced wage costs.

The average monthly incidence is lower for the selected sample (15.4 percent) than for the gross sample (20.0 percent). The average sick leave rates for these groups (weighed for both the duration of sick leave within the month and the length of the working week) are 9.9 percent and 5.4 percent, respectively. Thus, by selecting workers who were not ill on all working days of the previous month, a (small) group of workers, who are relatively prone to be absent and not likely to participate in the lottery, and constitute a major part of the total sick leave rate, is excluded from the sample. The remaining observable characteristics do not differ substantially between both samples: about 19 percent of the observations are female workers; about 32 percent of the workers have non-native parents; about 17 percent of the workers have a part-time job; the average age is 41.6 years and the average tenure is 12.7 years.<sup>5</sup> About 41 percent of the workers are employed in the smaller plant.<sup>6</sup>

Table 1 also compares the sick leave incidence and sick leave rates before and during the lottery. We find only the sick leave rate to have changed significantly after the introduction of the lottery, both for the gross and the selected sample. This is somewhat surprising, as participation in the lottery depends on the past incidence of absenteeism rather than the rate. We also compared the

---

<sup>5</sup> As we know the date of birth as well as the (exact) date of entry at the firm of individual workers, we know exactly the age and tenure.

<sup>6</sup> The summary statistics at the worker level (481) observations do not differ significantly from the dataset of 9,637 monthly observations, except for location.

composition of workers who were hired before and after announcement of the lottery. Although the introduction of an incentive system could lead to an additional inflow of less risk-averse workers (Lazear, 2000), it seems not to be very likely that it happened in the lottery of this firm, in which the lottery bonus was non-recurrent. Indeed, for the averages of the hirees' observables, we do not find significant differences between both periods.

<Table 2>

Table 2 presents the statistics for the group of lottery winners. When comparing this information on sick leaves with the information in Table 1, the incidence (7.0 percent) and monthly duration (1.9 percent) of sick leaves are considerably smaller. This is not surprising, as (on average) lottery winners participated in more lotteries than non-winners, and one of the conditions for lottery participation is not having reported ill for three consecutive months. Thus, winners are likely to be healthier than non-winners. Furthermore, we find both the incidence and the rate of sick leave to have decreased after the start of the lottery, suggesting that – all else equal – the lottery incentive was effective. The remaining variables show no substantial differences between the two samples, with the exception of the location of the plants. This stems from the fact that the stratification of the lottery across the plants (3:4) does not exactly correspond to the fraction of workers in both plants (2:3).

<Table 3>

We have information both on the worker's department within the firm and on his job title. Based on information from the personnel department, we distinguished the workers into 22 departments and the job titles into 12 job levels, corresponding to salary scales. It should be noticed that we do not observe the (exact) wages that are earned. Table 3 gives the fraction of workers in each of the salary scales, as well as the minimum and maximum monthly gross salary. A substantial fraction of the workers is concentrated in the lower-ranked job levels. More particularly, about 80 percent of the workers have a function below job level 8. The maximum gross monthly salary varies between 1,868 euros and 3,582 euros, but for the higher-ranked job level information on the salary is lacking. Thus, in terms of a percentage of the net wage, winning the lottery results in an additional

monthly net pay of 4 to 6 percent.

*<Figures 1,2>*

Figure 1 presents the monthly rate of sick leave over the period July 2001 – July 2003. The rate is weighed for the duration of sick leave during the month as well as for the length of the working week. As from September 2001, we observe a downward trend in the monthly rate of sick leave. It is however unclear if, and to what extent, this trend can be attributed to the lottery, which started in June 2002. Figure 2 portrays the development of the average monthly incidence of sick leave, which we will use as dependent variable in the regression analysis. Again, we observe a downward trend in the rate of sick leave.

*<Figure 3>*

Next, we focus on the workers' eligibility for the lottery, for the period June 2002 – July 2003. Figure 3 distinguishes between three groups of workers: workers who have won a previous lottery and therefore are no longer eligible to the lottery (lower area); workers who participate in the lottery as they had neither been on sick leave in the previous three months nor have won a lottery (upper area); and workers who had been on sick leave on any working day of the past three months and therefore are not eligible for the lottery (middle area). Figure 3 shows that by the end of July 2003 about 20 percent of the work force have won one of the lotteries. The probability of being on sick leave in the past three months varies between 20 percent and 30 percent during the lotteries. Finally, conditional on participating in the lottery, the monthly probability of winning increases from 2.7 percent (June 2002) to 3.4 percent (July 2003). The reason for this is that the number of eligible workers declines over time, whereas the number of (monthly) winners remains constant. As a result, the stimulating effect from the lottery scheme increases somewhat for workers who have not yet won the lottery.<sup>7</sup>

*<Figures 4,5>*

Figures 4 and 5 consider participation in the lotteries in more detail. At the one extreme, if all lottery winners had participated in all lotteries (before they actually won), we would observe a downward sloping fraction of lottery winners

---

<sup>7</sup> Obviously, this time effect would have been more substantial if there were no inflow of new workers into the group of eligible participants for the lottery.

over the number of lotteries. Figure 4 clearly does not support this hypothesis – winners did not participate in all lotteries. Furthermore, for the non-winners in Figure 5, we find a substantial fraction of workers (35%) who have participated in all lotteries. This stresses the importance of the heterogeneity of workers in the sample: some of them have a very low sick leave probability and are likely to participate in (almost) all lotteries; others participate in the lottery occasionally.

#### 4. Empirical model

In this section, we discuss how the incentive of the lottery can be incorporated in an empirical model of the monthly incidence of sick leave. For ease of exposition, we start by investigating the consequences of the lottery bonus for the sick leave decision within a month. For the moment, we assume that in the coming months there is a lottery and that the worker has not yet won any of the lotteries. In this setting, we have modeled the worker's sick leave decision as a standard reservation strategy (in the appendix of this paper). For an algebraic and detailed exposition of the lottery incentive effect, we refer to the appendix of this paper

Due to the introduction of the lottery the reservation strategy of a worker changes over time, causing the incidence of sick leave to change over time as well. Within a particular month, we may expect the lottery incentive of an individual worker to increase. As time proceeds and the worker has not become sick, the more likely the prospect of participating in the lottery will become – thus strengthening the incentive. At the same time, the fraction of workers that in fact has reported ill increases, thus decreasing the overall lottery incentive. If workers are risk-neutral, it can be shown that both effects are equal. Hence, the incentive effect will remain constant within a particular month.

The lottery incentive over three consecutive months is more complex, since absenteeism in the current calendar month  $m$  has consequences for participation in the lotteries of the next three months. Each of these lotteries may give a stimulus to the worker, although future incentives may have a lower value to the worker, which is picked up by a monthly discount rate  $\rho$ . In what follows, subscript  $i$  refers to the  $i$ -th worker ( $i = 1, \dots, N$ ) and subscript  $m$  to the  $m$ -th calendar month ( $m = 1, \dots, M$ ).

Furthermore, we introduce an indicator for the monthly incidence of absenteeism  $S_{i,m}$  which is 1 if the  $i$ -th worker has been on sick leave during the  $m$ -th month and zero if otherwise. Given that there will be a lottery in the coming months and the worker has not won any of the lotteries yet, three future lotteries may affect the current decision not to report sick in the  $m$ -th month. The compound of these three incentives is represented by  $I^\rho$ , which measures the incentive of (future) lotteries:

$$(1) \quad I_{i,m}^\rho = (1 - S_{i,m-1})(1 - S_{i,m-2}) + \rho(1 - S_{i,m-1}) + \rho^2 \quad i = 1, \dots, N; \quad m = 1, \dots, M$$

Equation (1) shows that there are three impetuses to abstain from absence in the current month  $m$ . First, when he has not been on sick leave in the months  $m-1$  and  $m-2$ , the worker may participate in the next lottery at the beginning of month  $m+1$ .<sup>8</sup> In addition, when he has not been on sick leave in  $m-1$ , he may participate in the lottery of month  $m+2$  (discounted by  $\rho$ ). Finally, the worker may participate in the lottery of month  $m+3$  (discounted by  $\rho^2$ ), irrespective of sick leave behavior in the past months.<sup>9</sup>

In equation (1),  $I^\rho$  explicitly depends on the discount factor  $\rho$ , which gives the value the worker attaches to the two consecutive lotteries following the upcoming lottery. For  $\rho = 0$ , the worker appreciates the upcoming lottery only, which could be due to myopic behavior. Experimental studies often stress the importance of hyperbolic discounting and awareness of the incentive (Frederick et al., 2002). This may also be relevant within the context of the lottery we investigate. That is, workers may not be (fully) aware of the consequences of their absenteeism on consecutive lotteries, or they are interested in the upcoming lottery only. We return to these issues when discussing the estimates of the model. For  $\rho = 1$ , the worker values the three lotteries equally. Discounting may occur as a result of time preferences, as well as with respect to the probability of reporting sick in the next (relevant) months. From this perspective, we may expect  $\rho$  to be higher for

---

<sup>8</sup> As the decision is taken for the whole month, and the lottery is held on one of the first working days of the next month  $m + 1$ , we do not need to discount the revenue.

<sup>9</sup> Equation (1) is used from July 2002 onwards. In the starting-up period of the lottery the incentive differed, as it referred to sick leave behavior after the announcement date in the beginning of May 2002 (see Footnote 2). Thus, in May 2002, the incentive is  $I_{i,m}^\rho = 1 + \rho + \rho^2$ ; in June 2002, it becomes  $I_{i,m}^\rho = (1 - S_{i,m-1}) + \rho + \rho^2$ . Our empirical results do not change substantially when we exclude all observations of May 2002 and June 2002 from the sample.

healthy workers.

In addition to the incentive  $I^\rho$  in equation (1), two conditions have to be met for participation in the lottery. The first condition is that a lottery is held in the coming month, which pertains to the observations from May 2002 onwards. The second condition for lottery participation is that the worker has not won one of the previous monthly lotteries yet. By implication, equation (1) becomes invalid when a worker wins the lottery.<sup>10</sup> Conditional on participation in the lottery, winning the lottery is random. Consequently, this source of variation can be considered exogenous. All together, the incentive variable we will use in the empirical analysis equals:

$$(2) \quad Z_{i,m}^\rho = I_{i,m}^\rho * L_m * E_{i,m} \quad i = 1, \dots, N; \quad m = 1, \dots, M$$

where the indicator function  $L_m$  is 1 when a monthly lottery is held, (in  $m+1$ ), and zero if otherwise (thus,  $L_m$  is 1 from May 2002 onwards).  $E_{i,m}$  is 1 if the  $i$ -th worker has not won any lottery up to the  $m$ -th month, and zero in all other cases including the month in which the lottery was won.

In general, the monthly incidence of sick leave is

$$(3) \quad S_{i,m} = \alpha Z_{i,m}^\rho + f(\theta_i, X_i, T_m) + \varepsilon_{i,m} \quad i = 1, \dots, N; \quad m = 1, \dots, M$$

with  $\alpha \leq 0$ . Besides the influence of the monthly lottery, the monthly incidence depends on a function  $f(\cdot)$  of unobserved worker-specific effects ( $\theta_i$ ), a vector of observed characteristics of the worker ( $X_i$ ), and a vector of time-varying information ( $T_m$ ).  $\varepsilon_{i,m}$  is an i.i.d. random error term.

Basically, the worker-specific effect,  $\theta_i$ , allows for differences in general health status across workers. However, there may be a severe, temporal drop in a worker's health condition, which results in a prolonged term of sick leave. Individual workers who are incapable of working through a severe illness cannot respond to the incentive during such a period at all. We assume that workers in protracted illness are no part of the population of workers who are capable to respond to the monthly lottery. This assumption results in the following sampling

---

<sup>10</sup> To be precise: there is no incentive in the month in which a worker wins a lottery, as the lottery is held on the first working day of the month as a consequence of which the incentive is absent in the

criterion: those observations  $S_{i,m}$  for which worker  $i$  has been on sick leave on all working days of the previous calendar month  $m-1$  have been excluded from the sample. It implies that for all of the observations of the sample, showing up on the work floor is within the bounds of possibility of individual workers.

To estimate equation (3), we follow the Probit specification, in which the monthly incidence of sick leave  $S$  depends on:<sup>11</sup>

$$(4) \quad S_{i,m} = \alpha Z_{i,m}^\rho + \beta \overline{Z}_i^\rho + \gamma' X_i + \delta' T_m + \varepsilon_{i,m} \quad i = 1, \dots, N; \quad m = 1, \dots, M$$

where  $X$  is a vector of time-invariant,<sup>12</sup> observable worker characteristics, which include age, tenure, dummies for gender, part-time, non-native parents and plant, as well as dummies for both the function level and the department of the worker.  $T$  is a vector of calendar month dummies (24 dummies).  $\varepsilon$  is a Normally distributed error term. The parameter values of the vectors  $\gamma$  and  $\delta$  as well as the scalars  $\alpha$  and  $\beta$  register the effect of  $X$ ,  $T_m$ ,  $Z^\rho$  and  $\overline{Z}^\rho$ , respectively. As the incentive variable  $Z^\rho$  is the only independent variable that varies both across workers and over time, we included its average per worker  $\overline{Z}^\rho$  (averaged over all months in which the worker is observed) to control for the unobserved worker effect.<sup>13</sup> The intuition behind this approach is that  $\beta$  is estimated by changing  $Z^\rho$ , but holding the time average constant.<sup>14</sup> As a result, we arrive at the Random Effects Probit model. Maximum likelihood gives a consistent estimator of our parameter of interest,  $\beta$ , as the correlation of  $\overline{Z}^\rho$  with  $\varepsilon$  does not seriously affect exogeneity – that is,  $M = 25$  months is fairly large.

The incentive  $Z^\rho$  variable depends on the discount rate  $\rho$ . In order to estimate the model, we will apply a grid-search procedure, in which we opt for the value of  $\rho$  that gives the lowest absolute value of the log likelihood function.

remainder of this month.

<sup>11</sup> We estimate the model at the monthly level, as it corresponds to the frequency of the lotteries, which is sufficient to obtain a consistent estimate of the incentive effect.

<sup>12</sup> Of course, the age and tenure of workers increase slightly over the period of investigation (25 months).

<sup>13</sup> In order to obtain standard errors of  $\rho$ , we should have estimated the model with  $\rho$  as a separate variable. Estimation of such a model is however rather cumbersome, as this would involve nonlinear constraints in the model (see equation (1)). Moreover, as we will show in the following, there is a risk of obtaining corner-solutions ( $\rho \geq 0$ ).

Moreover, the estimation results for the various model versions can be used as a robustness check. At the one extreme, with  $\rho = 0$ , workers are stimulated only by the next lottery. At the other extreme, with  $\rho = 1$ , workers fully take into account the two subsequent lotteries as well.

<Table 4, 5>

## 5. Estimation results

### 5.1 The benchmark model

This sub-section discusses the estimates of equation (4) for those observations for which the worker was not on sick leave on all of the working days of the previous calendar month. For various values of  $\rho$ , we calculated the corresponding value of the log likelihood. Table 4 gives the outcomes for both the net sample and the sample of winners. Note that the coefficient estimates in the table are expressed as marginal effects (in percentage points). Apparently, the lowest absolute value of the log likelihood is for  $\rho = 0$ . At this extreme, workers are driven by the first upcoming lottery only, and myopic in their valuation of future lotteries. Two explanations for this result are that workers were not (fully) aware of the consequence of current absenteeism for future lotteries, or that they were only interested in the first upcoming lottery due to hyperbolic discounting. Still, careful interpretation is needed when analysing our grid-search estimation results. In particular, we can reject the null hypothesis that workers value all upcoming lotteries equally ( $\rho = 1$ ), but we cannot conclude from this that  $\rho$  equals zero.

The result that discounting is important holds for both the net sample and the sample of lottery winners. Lottery winners, who are more likely to be in good health, are also more likely to participate in future lotteries. Therefore, we may expect this group to have a higher rate  $\rho$ . Our results suggest that there is no evidence for this, but the power of our grid search test is limited here.

Table 5 presents the estimated parameters for  $\rho = 0$ , i.e. the model where workers are driven by the first upcoming lottery only. Most of the coefficient

---

<sup>14</sup> This model is often referred to as Chamberlain's Random Effects model (Chamberlain, 1982).



estimates of  $\gamma$  are insignificant. Only the dummies describing calendar time effects and the dummies for the various departments of the firm and for the function levels (salary scale) differ significantly from zero.

Generally, we find the estimated coefficient on the lottery incentive to be significant different from zero and economically substantial. For the net sample, participation in the lottery induces workers to report sick less frequently, with the effect amounting to a 2.1 percentage point reduction in the monthly sickness incidence. Their average sickness incidence is 15.4 percent (see Table 1). Remarkably, we find the estimated coefficient on the lottery incentive to be higher for the sub-sample of winners than for the net sample. For the winners, the estimated marginal effect equals 4.1 percent point. This is substantial, in particular compared with the average sickness incidence in the sample of winners of 8.8 percent (see Table 2). It suggests that the lottery winners responded strongly to the lottery incentive. As we have argued earlier, workers who are in good health are likely to participate in the lottery more frequently, making it more likely that they will win a lottery. Furthermore, healthy workers who report sick are likely to have short sickness durations, e.g. one or two days. Thus, the lottery incentive induces these workers to decide whether they should not report sick at all. When they do so, it causes a sizeable reduction in the incidence of sickness rather than in the rate of sickness.

The estimates results suggest that, from the perspective of the firm, the lottery was highly beneficial. For an average monthly salary of 1800 euros, and a 50% participation rate in the lottery, the estimated monthly amount of benefits in terms of sickness pay reduction equals 5,760 euro.<sup>15</sup> Hence, the benefits dwarf the monthly costs of the lottery, which are a mere 525 euro. In the first lotteries, workers may have overestimated the probability of winning the lottery. However, this effect is not likely to persist over time, and – given the size of the effect – can

---

<sup>15</sup> The average estimated lottery effect on the sickness rate equals 1.6%-point (this can be derived from the Tobit model, as we will show in Sub-section 5.2); the average number of workers in the sample was about 400, and on average about 50% of these workers met the conditions for participation in the first upcoming lottery (see Figure 3). With an average wage of 1,800 euros, the savings on sickness absenteeism costs equal  $400 \times 1.6\% \times 50\% \times 1,800 \text{ euros} = 5,760 \text{ euros}$  per month.

only be a modest and partial explanation. Non-monetary aspects are more likely to drive our results. Workers may be intrinsically motivated<sup>16</sup> to participate in the lottery, just because of fun. Furthermore, winning the lottery entails more than receiving a bonus. Lottery winners are publicly announced, thus acting as a signal that these workers are less likely to report sick than their colleagues. Workers may also like the idea of earning more than their colleagues, rather than the exact amount of money that is involved.<sup>17</sup>

## 5.2 Robustness checks

So far, we have made various assumptions regarding the specification of equation (3), in particular with respect to the response mechanism of individual workers to the lottery scheme. This sub-section investigates the consequences of relaxing or changing these assumptions.

To start with, we re-estimate the model following the (conditional) Logit specification with fixed worker effects, instead of the Random Effects Probit model.<sup>18</sup> It results in an insignificant estimate of the incentive effect for the full sample, and an estimated coefficient of about  $-0.060$  (which is statistically significant at the 5%-level) for the sub-sample of winners.<sup>19</sup> However, this estimation procedure may be hampered by the fact that the model cannot be estimated using information of healthy workers who have not been on sick leave for the whole sampling period.

---

<sup>16</sup> Another explanation is that extrinsic motivation has crowded out the intrinsic motivation of workers not to report sick. In particular, workers who are not or no longer eligible for the upcoming lottery may be triggered to report sick more often. This causes the incentive effect – which is based on the comparison of those who are eligible with those who are not eligible for the lottery – to be overestimated. A natural way to test for this effect is to compare absenteeism behavior of workers before the lottery, during the lottery, and after winning the lottery (see Sub-section 5.2).

<sup>17</sup> In addition, we find no difference in the effect of the incentive variable for various specific groups of workers. We added in equation (4) an interaction term in the incentive variable and one of the following dummy variables (as well as the average per worker of this cross term): a dummy for older workers, for workers with long tenure, female workers, part-time workers, workers with non-native parents, and workers with high-function levels. We estimated this equation for each of these cross-terms separately. For none of these specifications, the estimated coefficient on the cross-term is significantly different from zero.

<sup>18</sup> Within the context of the Conditional Logit model, we should be aware that the number of explanatory variables decreases considerably; parameter estimates can only be identified for those variables that vary over time. We included 24 monthly dummies and the lottery incentive.

<sup>19</sup> The Conditional Logit estimates are based on a sub-sample of 344 workers (65 lottery winners)

As a second robustness check, we investigate sick leave behavior of workers *after* they won the lottery. As we have argued earlier, we can distinguish between two meaningful types of information that identify the lottery incentive effect. Hence, the incentive effect is over-identified. First, there is variation in past sickness behavior, which determines eligibility for the lottery. Second, workers are no longer eligible for the lottery after they have won the lottery, which also results in variation in the incentive. The over-identification of the incentive effect enables us to test for the presence of persistency in sickness behavior. Suppose that lottery winners show a (permanently) reduced incidence of sickness, then the incentive estimate that is based on the second type of information – comparing lottery winners before and after winning the lottery – will be underestimated. We test persistency behavior by including a ‘persistency dummy’ as an additional variable in the model – that is, a dummy describing the effect after winning the lottery. We rewrite equation (4) as

$$(5) \quad S_{i,m} = \alpha_1 Z_{i,m}^p + \beta_1 \overline{Z_i^p} + \alpha_2 W_{i,m} + \beta_2 \overline{W_i} + \gamma' X_i + \delta' T_m + \varepsilon_{i,m}$$

with

$$W_{im} = (1 - E_{im})$$

and

$$i = 1, \dots, N_w; \quad m = 1, \dots, M$$

where  $N_w$  is the number of lottery winners over the full sampling period. The ‘persistency’ dummy variable  $W_{im}$  is one once a worker has won a lottery (so it becomes one in the month of winning) and zero elsewhere. If persistency behavior is important, we may expect  $\alpha_2$  to be smaller than zero, indicating that the incentive effect persists and sickness incidence does not fall back to the level without lottery incentive.

<Table 6>

Table 6 presents the estimation results of equation (5), based on the sub-sample of workers who have won the lottery.<sup>20</sup> The coefficient on the persistency

---

who have been on sick leave in at least one of the 25 months of the sampling period.

<sup>20</sup> Obviously, estimation of the specification with persistency dummy for a sample of winners and non-winners is likely to result in biased estimates. Therefore, the persistency effect is estimated

dummy  $W_{in}$  is insignificant. Furthermore, the estimated coefficient on the incentive – which is now based on the first type of information (variation in the past sickness behavior of workers) only – does not change substantially. It suggests that persistency effects are not important.

As a third robustness check, we concentrate on the impact of the lottery incentive on the rate of sickness vis-à-vis the incidence. In the benchmark model (defined in equation (4)), we concentrated on the impact on the monthly incidence of sickness only – the (implicit) assumption being that the distribution of sickness rates has remained unaffected. However, it may be that the effect of the incentive has led to a reduction of short sickness durations only – causing the sickness rate to decrease only slightly compared to the sickness incidence.

In order to get a more complete picture of the effect on the sickness rate, we re-estimate the model using a Tobit specification, in which the dependent variable is the fraction of working days of the month in which the worker is on sick leave. First, it helps us to obtain the marginal effect of the incentive variable, on both the incidence and the rate of sickness. Second, the comparison between the estimation results of the Probit and the Tobit model provides us with a natural test on sensitivity of worker groups on the incentive effect. For the Probit model, the incentive coefficient estimate is identified from the (discrete) data on sickness incidence only. For the Tobit model, identification follows from data on both the sickness incidence and the (conditional) sickness rate. Now, if the incentive effect is confined to (the density of) low sickness rates only, the estimated marginal effects of both models will not coincide. In particular, the coefficient estimate of the marginal effect of the incentive in the Probit model would exceed that of the Tobit model.

*<Tables 7 and 8>*

Tables 7 and 8 present the parameter estimates of the Tobit model for the sickness incidence and the sickness rate of the net sample and the sample of winners, respectively. Note that the coefficient estimates are again expressed as marginal effects. We find the estimated impact of the lottery incentive on the

---

using the information on winners only, whereas the effect of the lottery is based on information of both winners and non-winners.

incidence and the rate of sickness for the net sample to equal 3.5 percentage points and 1.6 percentage points, respectively. Similarly, the estimated effects are 5.1 percentage points and 1.8 percentage points for the sub-sample of winners. Thus, there is no significant difference between the incentive effect on the rate of sickness for the net sample and for the sub-sample of winners.

The estimated coefficients on the sickness incidence exceed those in the Probit model – the difference being insignificant for both samples. From this, we conclude that the incentive effect has equally affected the incidences of sickness spells of short and long durations.

## **6. Conclusion**

We have investigated the effect of a unique monthly lottery established by a large Dutch manufacturer with the intention to stimulate work floor attendance. So far, the literature on sickness absence behavior has mostly focused on negative incentives – that is, the effect of wage decreases upon absence, or the future consequences of absence for the labor market position of workers. Our study contributes to this literature by analyzing the effect of a bonus system on individual absenteeism. In the lottery we investigated the incentive bonus system was directly related to individual absence behavior in the past three months. Furthermore, our analysis is related to the empirical literature on the effect of incentives on productivity (initiated by Lazear (2000)). In general, these studies show that incentive pay has a positive effect on individual productivity. Loosely speaking, the lottery bonus may be interpreted as a kind of incentive pay and absenteeism as a measure of productivity.

We have estimated a semi-structural model; i.e. we specified the incentive effect (which depends on past absenteeism) to be highest for those workers who are (still) eligible for all of the three upcoming lotteries. According to our estimation results, the lottery leads to a substantial decrease in the incidence of sickness of 3.5 percentage point. This corresponds to a reduction in the rate of sickness of 1.6 percentage point. Interestingly, it seems that workers are mainly driven by the first upcoming lottery. When looking at the effect of the lottery on the incidence of sickness, we find it to be particularly strong for those workers who have won the

lottery during the period of investigation. Presumably, these workers were in better health, making it more likely for them to win the bonus. In addition, healthy workers who report sick are likely to have short sickness durations, of e.g. one or two days. Thus, the lottery incentive induces these workers to decide not to report sick at all, causing a sizeable reduction in the incidence of sickness, rather than the rate of sickness.

Our results suggest that the lottery was highly beneficial to the firm. It therefore seems that non-monetary aspects to workers are important in explaining the substantial incentive effect. One of these aspects may be the public announcement of the winners among the colleagues. Furthermore, the idea of earning more than colleagues may be more important to workers than the exact amount of money involved. However, we may expect that the effect of the lottery will diminish after some time, since we find that winning workers resume their (higher) rate of absence prior to the lottery. It seems that this incentive system cannot alter workers' identities (Akerlof and Kranton, 2005). Probably, the incentive system might have a longer lasting effect on absenteeism if the lottery winners remain eligible for future lotteries or if the criteria for participation in the lottery were loosened, by e.g. extending participation to workers who do not report sick for two, instead of three months. However, from our results we cannot conclude that such changes would reduce absenteeism to a greater extent. Clearly, the number of workers who are stimulated not to report sick will increase, but in turn this might reduce the incentive effect of the lottery as a signaling device.

## References

- Akerlof, G.A. and R.E. Kranton (2005), "Identity and the Economics of Organizations", *Journal of Economic Perspectives*, 19 (1), 9-32.
- Allen, S.G. (1981), "An Empirical Model of Work Attendance", *Review of Economics and Statistics*, 63 (1), 77-87.
- Arai, M. and P.S. Thoursie (2005), "Incentives and Selection in Cyclical Absenteeism", *Labour Economics*, Vol. 12, 269-280.
- Barmby, T.A., M.G. Ercolani, and J.G. Treble (2002), "Sickness Absence: An International Comparison", *Economic Journal*, 112 (480), F315-F331.
- Barmby, T.A., Chr. Orme and J.G. Treble (1991), "Worker Absenteeism: An Analysis Using Microdata", *Economic Journal*, 101 (405), 214-229.
- Barmby, T., Chr. Orme and J. Treble (1995), "Worker Absence Histories: A Panel Data Study", *Labour Economics*, 2 (1), 53-65.
- Besseling, J.J.M., M.J. van Gent and E. Brouwer (1999), "Van TAV tot REA: bedrijfsbeleid en verzuim", Ministerie van Sociale Zaken en Werkgelegenheid, Elseviers bedrijfsinformatie, 's Gravenhage.
- Brown, S., Fakhfakh, F. and J.G. Sessions (1999), "Absenteeism and Employee Sharing: An Empirical Analysis based on French Panel Data, 1981-1991", *Industrial and Labor Relations Review*, 52 (2), 234-251.
- Chamberlain, G. (1982), "Multivariate Regression Models for Panel Data", *Journal of Econometrics*, 18 (1), 5-46.
- Doherty, N.A. (1979), "National Insurance and Absence from Work", *Economic Journal*, 89 (353), 50-65.
- Drago, R. and M. Wooden (1992), "The Determinants of Labor Absence: Economic Factors and Workgroup Norms across Countries", *Industrial and Labor Relations Review*, 45 (4), 764-778.
- Dunn, L.F. and S.A. Youngblood (1986), "Absenteeism as a Mechanism for Approaching an Optimal Labor Market Equilibrium: An Empirical Study", *The Review of Economics and Statistics*, 68 (4), 668-674.
- Engellandt, A. and R.T. Riphahn (2004), "Incentive effects of Bonus Payments: Evidence from an International Company", IZA Discussion Paper, No. 1229.
- Einerhand, M.G.K., G. Knol, R. Prins and T.J. Veerman (1995), "Sickness and invalidity arrangements – Facts and figures from six European Countries", Ministry of Social Affairs and Employment, The Hague, The Netherlands.

Frederick, S., G. Loewenstein and T. O'Donoghue (2002), "Time Discounting and Time Preference: A Critical Review," *Journal of Economic Literature*, XL (2), 351-401.

Henrekson, M. and M. Persson (2004), "The Effects on Sick Leave of Changes in the Sickness Insurance System", *Journal of Labor Economics*, 22 (1), 87-113.

Ichino, A. and R.T. Riphahn (2004), "The Effect of Employment Protection on Worker Effort: A Comparison of Absenteeism During and After Probation," forthcoming in *Journal of the European Economic Association*.

Johansson, P. and M. Palme (1996) "Do Economic Incentives Affect Work Absence? Empirical Evidence Using Swedish Micro Data", *Journal of Public Economics*, 59 (2), 195-218.

Johansson, P. and M. Palme (2001) "Assessing the Effect of Public Policy on Worker Absenteeism", *Journal of Human Resources*, 37 (2), 381-409.

Jong, Ph. de and M. Lindeboom (2003), "Privatisation of sickness insurance: Evidence from the Netherlands", Paper presented at conference "Sickness absence: diagnoses and cures" (Economic Counsel of Sweden), October.

Kaivanto, K. (1997), "An Alternative Model of Pro-cyclical Absenteeism," *Economics Letters*, 54 (1), 29-34.

Kenyon, P. and P. Dawkins (1989), "A Time Series Analysis of Absence in Australia", *Review of Economics and Statistics*, 71 (2), 232-239.

Lazear, E. P. (2000), "Performance Pay and Productivity", *American Economic Review*, 90 (5), 1346-1361.

Leigh, J.P. (1985) "The Effects of Unemployment and the Business Cycle on Absenteeism", *Journal of Economics and Business*, 37 (2), 159-170.

Wilson, N. and M. J. Peel (1991), "The Impact on Absenteeism and Quits of Profit-Sharing and Other Forms of Employee Participation", *Industrial and Labor Relations Review*, 44 (3), 454-468.

Winkelmann, R. (1999), "Wages, Firm Size and Absenteeism", *Applied Economics Letters*, 6 (6), 337-341.

Winkler, D.R. (1980) "The Effect of Sick-Leave Policy on Teacher Absenteeism", *Industrial and Labor Relations Review*, 33 (2), 232-240.



## Appendix: Modeling the effect of the lottery incentive

In this appendix we analyze the effect of the lottery incentive on absenteeism. In particular, we are interested in the variables driving the effect of the lottery incentive, as well as its evolution within a particular month. We start by modeling the workers sick leave decision as a standard reservation strategy:

$$(I) S_d = I(h_d > H^*)$$

with  $d = 1, \dots, D$ .  $S_d$  represents the rate of sick leave of a worker at day  $d$ ,  $h_d$  the health status at  $d$ , and  $H^*$  the reservation value of  $h$  for which the worker is indifferent between sick leave and being at work.  $H^*$  is determined by the value of leisure of a particular worker, and the health status  $h$  is drawn from a distribution  $G$ .  $I$  indicates the event between parentheses. Under these assumptions, the probability of reporting sick at day  $d$  is equal to  $1 - G(H^*)$ .

When expressing  $H^*$  in terms of money utility value, the model can be extended with the bonus incentive. We denote the expected value of participating in the lottery as  $L$ . The possible prospect of a bonus lies at the end of the month,  $D$ , provided that the worker has not reported sick until  $d$ . Due to the introduction of the bonus, we may expect the reservation strategy of the worker to change over time. The dynamic programming problem of a worker – represented by the reservation value  $H^*(d)$  – who has not reported sick until time  $d$  can then be written as:

$$(II) \quad H^*(d) = H^* + L \exp\left\{\int_d^D [1 - G(H^*(s))] ds\right\}$$

with  $0 < d < D$ .

Equation (II) shows that the incentive effect of the prospect of a bonus increases over time, provided that sick leave has not occurred so far. The expected money value of the bonus is equal to the expected value of participation in the lottery  $L$ , discounted by the survival probability of not reporting sick until  $D$ . If however a worker decides to report sick at time  $d$ , we return to the model as in equation (I). For small values of  $L$ , the effect of the incentive can be approximated by a first order Taylor series expansion at  $L = 0$ :

$$(III) \quad \frac{\partial G(H^*(d) | L=0)}{\partial L} = g(H^*) \exp\{-(D-d)[1-G(H^*)]\}$$

This expression shows that the impact of the bonus incentive is determined by three variables. First, the model predicts the impact to depend on the density of  $G$  at  $H^*$ , just as any other variable that affects  $H^*$  and is constant over time. Second, the impact of the incentive increases over time. This is not surprising, as the prospect of participating in the lottery becomes more likely over time. Third, the model predicts this increase over time to interact with the ex ante sickness probability of workers. The intuition behind this result is that the incentive is stronger for healthy workers who are more likely to participate in the lottery.

Combining the above findings, the net impact of  $H^*$  on the bonus effect is ambiguous. On the one hand, for workers with a low sick leave rate (and small  $H^*$ ) the further downward potential of their sick leave rate is likely to be small – this can be labeled as the ‘density effect’. On the other hand, the incentive impact itself for this group is substantial – this is the effect that is due to a shift in the support.

So far, we have derived the impact for workers who have not reported sick until time  $d$ . In order to calculate the average (overall) impact for workers, we have to take into account that there is no further incentive for those workers who have reported sick. Thus, equation (III) is multiplied with the probability of not reporting sick until time  $d$ :

$$(IV) \quad \frac{\partial G(H^*(d) | L=0)}{\partial L} \exp\{-d[1-G(H^*)]\} = g(H^*) \exp\{-D[1-G(H^*)]\}$$

Expression (IV) shows that the overall impact of the bonus is constant within a particular month. This means that the incentive increases for workers who have not reported sick, but the size of this group decreases to the same extent.

**Table 1 – Summary statistics, all workers**

	<b>Gross sample</b>		<b>Selection of population <sup>a)</sup></b>	
	<b>Mean</b>	<b>Std.deviation of mean</b>	<b>Mean</b>	<b>Std.deviation of mean</b>
Monthly incidence of sick leave	0.200	0.004	0.154	0.004
– before the lottery	0.207	0.005	0.158	0.005
– during the lottery	0.194	0.005	0.151	0.005
Duration of sick leave (fraction of month)	0.099	0.003	0.054	0.002
– before the lottery	0.110	0.004	0.063	0.002
– during the lottery	0.090	0.003	0.047	0.002
Proportion of part-time workers	0.177	0.004	0.173	0.004
Proportion with non-native parents	0.326	0.005	0.321	0.005
Gender (woman = 1; man = 0)	0.197	0.004	0.193	0.004
Age (in years)	41.57	0.10	41.54	0.10
Tenure (in years)	12.76	0.10	12.66	0.11
Location (plant A = 1; plant B = 0)	0.408	0.01	0.411	0.01
Number of observations	9,637		9,088	

<sup>a)</sup> Excluded are observations of workers who were on sick leave on all working days of the previous month.

**Table 2 – Summary statistics, lottery winners**

	Gross sample		Selection of population <sup>a)</sup>	
	Mean	Std.deviation of mean	Mean	Std.deviation of mean
Monthly incidence of sick leave	0.104	0.007	0.088	0.006
– before the lottery	0.149	0.011	0.118	0.009
– during the lottery	0.070	0.007	0.066	0.007
Duration of sick leave (fraction of month)	0.043	0.004	0.029	0.003
– before the lottery	0.073	0.007	0.045	0.003
– during the lottery	0.020*	0.003	0.017*	0.002
Proportion of part-time workers	0.168	0.008	0.166	0.008
Proportion with non-native parents	0.357	0.01	0.350	0.01
Gender (woman = 1; man = 0)	0.16	0.008	0.160	0.008
Age (in years)	41.96	0.21	41.98	0.21
Tenure (in years)	12.96	0.20	12.94	0.21
Location (plant A = 1; plant B = 0)	0.483	0.01	0.481	0.01
Number of observations	2,158		2,116	

<sup>a)</sup> Excluded are observations of workers who were on sick leave on all working days of the previous month.

**Table 3 – Job level, worker percentages, and minimum and maximum wages (481 Workers)**

<b>Function level</b>	<b>Percentage of workers</b>	<b>Cumulative Percentage of workers</b>	<b>Minimum monthly gross salary (Euros) March 1<sup>st</sup>, 2002</b>	<b>Maximum monthly gross salary (Euros) March 1<sup>st</sup>, 2002</b>
1	3.74	3.74	1,616	1,868
2	4.99	8.73	1,676	1,948
3	20.79	29.52	1,735	2,013
4	17.67	47.19	1,813	2,130
5	12.27	59.46	1,896	2,319
6	14.97	74.43	1,995	2,467
7	6.24	80.67	2,120	2,658
8	8.73	89.40	2,275	2,907
10	0.83	90.23	2,698	3,582
11	6.44	96.67	a)	a)
12	2.91	99.58	a)	a)
13	0.42	100.0	a)	a)

a) Salary is unknown for the highest job levels.

**Table 4 – ML-estimates of parameters of  $Z^\rho$  in equation (4) (heteroscedasticity adjusted standard errors), for different values of  $\rho$ ; t-values between parentheses**

	$\rho = 0$	$\rho = 0.25$	$\rho = 0.5$	$\rho = 0.75$	$\rho = 1$
<b>Net sample</b>					
$Z^\rho$	-0.021	-0.021	-0.018	-0.013	-0.009
	(-2.04)	(-2.38)	(-2.42)	(-2.22)	(-1.90)
Log Likelihood (N=9,088)	-3503.0	-3529.8	-3557.4	-3578.0	-3591.6
<b>Winners</b>					
$Z^\rho$	-0.041	-0.041	-0.032	-0.024	-0.019
	(-2.77)	(-2.92)	(-3.00)	(-3.03)	(-3.03)
Log Likelihood (N=2,116)	-551.5	-552.2	-553.1	-553.9	-554.6

**Table 5 – ML-estimates of the benchmark model: equation (4) with  $\rho = 0$  ;  
(heteroscedasticity adjusted standard errors)**

Independent variable	Net sample		Winners	
	Marginal effect	t-value	Marginal effect	t-value
$Z^{\rho}, \rho = 0$	-0.021	(-2.04)	-0.041	(-2.77)
$\overline{Z^{\rho}}, \rho = 0$	-0.273	(-8.23)	-0.144	(-2.75)
Tenure	0.0004	(0.19)	0.002	(0.61)
Tenure-squared / 100	-0.004	(-0.62)	-0.010	(-0.85)
Age	0.004	(0.76)	0.008	(1.31)
Age-squared / 100	-0.005	(-0.88)	-0.010	(-1.27)
Dummy gender (woman=1, man = 0)	0.028	(1.68)	0.009	(0.31)
Dummy part-time	-0.005	(-0.42)	-0.031	(-1.69)
Dummy non-native parents	-0.018	(-1.57)	0.026	(1.51)
Dummy plant	0.005	(0.20)	-0.065	(-2.19)
Calendar month dummies	$\chi^2_{(24)} = 194.9$		$\chi^2_{(24)} = 55.1$	
Function level dummies	$\chi^2_{(11)} = 47.1$		$\chi^2_{(8)} = 20.6$	
Department dummies	$\chi^2_{(21)} = 108.7$		$\chi^2_{(18)} = 39.4$	
Number of independent variables	66		60	
Pseudo R-squared	0.104		0.128	
Number of observations	9,088		2,116	
Log likelihood	-3503.0		-551.5	

**Table 6 – ML-estimates of equation (5), with  $\rho = 0$ , measuring persistency effects for winners (N=2,116); (heteroscedasticity adjusted standard errors)**

Independent variable	Marginal effect	t-value
$Z^{\rho}, \rho = 0$	-0.045	(-2.00)
$\overline{Z^{\rho}}, \rho = 0$	-0.267	(-2.51)
$W$	-0.008	(-0.31)
$\overline{W}$	-0.133	(-1.38)
Tenure	0.002	(0.68)
Tenure-squared / 100	-0.010	(-0.90)
Age	0.009	(1.50)
Age-squared / 100	-0.010	(-1.42)
Dummy gender (woman=1, man = 0)	0.007	(0.30)
Dummy part-time	-0.026	(-1.50)
Dummy non-native parents	0.028	(1.57)
Dummy plant	-0.062	(-2.17)
Calendar month dummies	$\chi^2_{(24)} = 45.8$	
Function level dummies	$\chi^2_{(8)} = 17.4$	
Department dummies	$\chi^2_{(18)} = 30.7$	
Number of independent variables	62	
Pseudo R-squared	0.130	
Number of observations	2,116	
Log likelihood	-550.2	



**Table 7 – ML-estimates of the Tobit model, dependent: (number of working days of sick leave in month)/ (total number of working days in month); net sample**

Independent variable	Effect on Sickness Incidence		Effect on Sickness Rate	
	Marginal effect	t-value	Marginal effect	t-value
$Z^{\rho}, \rho = 0$	-0.035	(-3.78)	-0.016	(-3.78)
$\overline{Z^{\rho}}, \rho = 0$	-0.228	(-11.66)	-0.103	(-11.58)
Tenure	0.002	(1.70)	0.001	(1.70)
Tenure-squared / 100	-0.008	(-2.11)	-0.004	(-2.11)
Age	0.004	(1.49)	0.002	(1.49)
Age-squared / 100	-0.006	(-1.64)	-0.003	(-1.64)
Dummy gender (woman=1, man = 0)	0.028	(2.83)	0.013	(2.83)
Dummy part-time	-0.004	(-0.34)	-0.002	(-0.34)
Dummy non-native parents	-0.009	(-1.19)	-0.004	(-1.18)
Dummy plant	0.007	(0.38)	0.003	(0.38)
Calendar month dummies		$\chi^2_{(24)} = 149.9$		
Function level dummies		$\chi^2_{(11)} = 60.6$		
Department dummies		$\chi^2_{(21)} = 101.7$		
Number of independent variables		66		
Pseudo R-squared		0.096		
Number of observations		9,088		
Log likelihood		-3675.6		

**Table 8 – ML-estimates of the Tobit model, dependent: (number of working days of sick leave in month)/ (total number of working days in month); sample of winners**

Independent variable	Effect on Sickness Incidence		Effect on Sickness Rate	
	Marginal effect	t-value	Marginal effect	t-value
$Z^{\rho}, \rho = 0$	-0.051	(-2.65)	-0.018	(-2.64)
$\overline{Z^{\rho}}, \rho = 0$	-0.120	(-2.70)	-0.042	(-2.67)
Tenure	0.003	(0.84)	0.001	(0.94)
Tenure-squared / 100	-0.011	(-1.15)	-0.004	(-1.15)
Age	0.009	(1.69)	0.003	(1.69)
Age-squared / 100	-0.010	(-1.66)	-0.004	(-1.66)
Dummy gender (woman=1, man = 0)	0.010	(0.45)	0.004	(0.45)
Dummy part-time	-0.028	(-1.31)	-0.010	(-1.31)
Dummy non-native parents	0.033	(1.93)	0.012	(1.93)
Dummy plant	-0.054	(-2.23)	-0.019	(-2.22)
Calendar month dummies		$\chi^2_{(24)} = 56.0$		
Function level dummies		$\chi^2_{(8)} = 20.9$		
Department dummies		$\chi^2_{(18)} = 50.3$		
Number of independent variables		50		
Pseudo R-squared		0.135		
Number of observations		2,116		
Log likelihood		-546.9		

Figure 1: Sick leave rate in gross sample  
July 2001 - July 2003

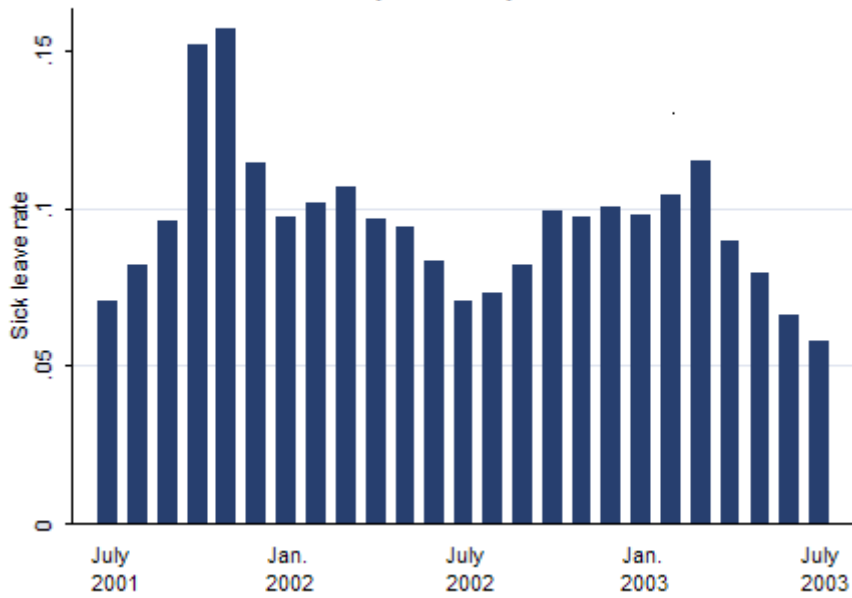


Figure 2: Sick leave incidence in gross sample  
July 2001 - July 2003

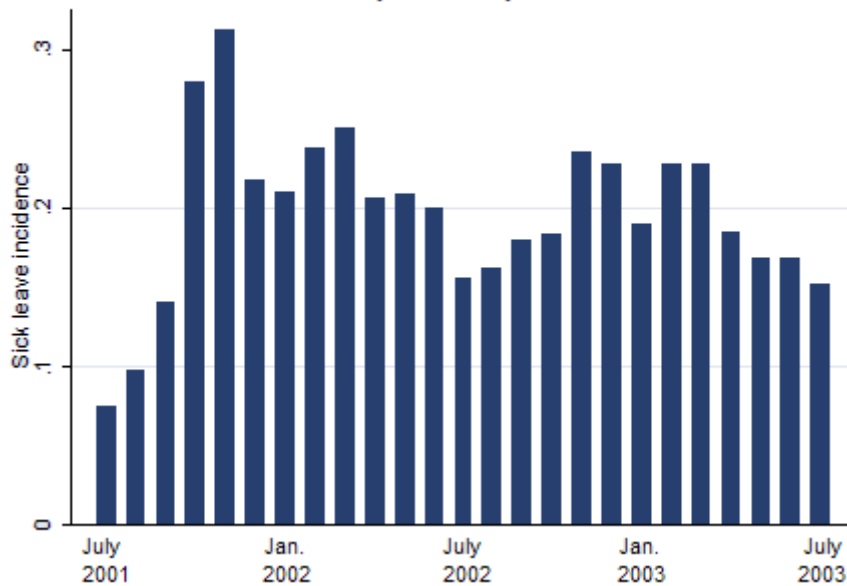


Figure 3: Fractions of participants in the lottery  
June 2002 - July 2003

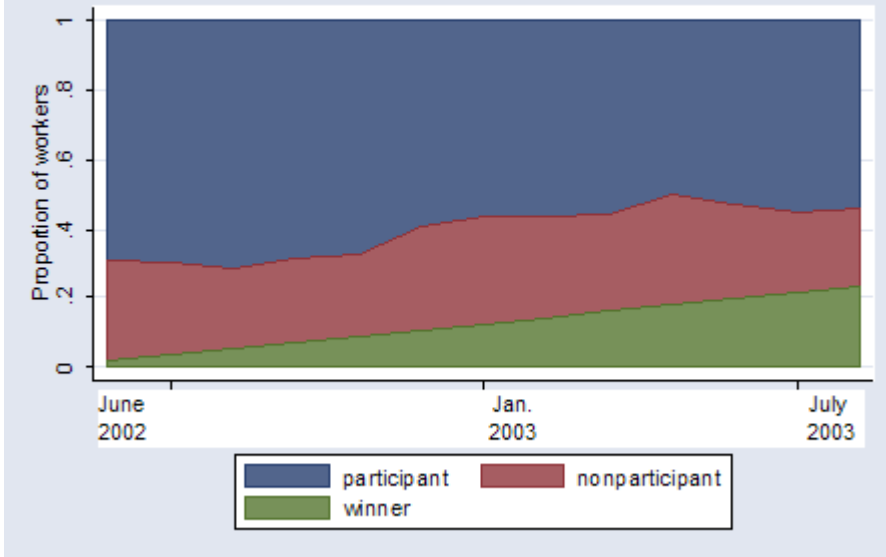


Figure 4: Number of lottery participations of lottery winners  
June 2002 - July 2003

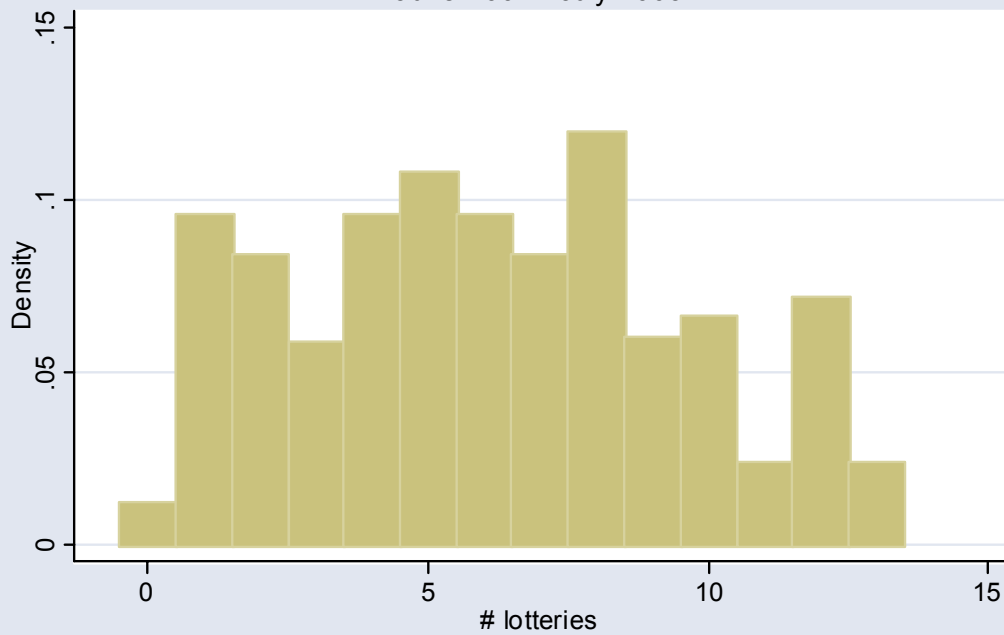


Figure 5: Number of lottery participations of non-winners  
June 2002 - July 2003

