University of Southern Denmark

COHERE - Centre of Health Economics Research, Department of Business and Economics
Discussion Papers, No. 2014:2
ISSN: 2246-3097

The Economics of Presenteeism: A discrete choice & count model framework

by
Kjeld Møller Pedersen[1]
Kristian Skagen[2]
[1,2]University of Southern Denmark

COHERE, Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55
DK-5230 Odense M
Denmark
www.cohere.dk

Abstract

There are three levels in this paper: A search for economic theories about presenteeism, a search for appropriate econometric approaches, and finally empirical results based on a unique Danish cross sectional data set.

There are two economic approaches to presenteeism: 1. Productivity losses and 2. labor supply. The first is part of the indirect cost component in cost-of-illness studies and economic evaluation. There are two core questions in the productivity loss literature: Measurement of productivity losses (‘how much’) which has dominated the research agenda and valuation of incurred productivity losses (monetary value). Few economists have addressed the valuation issue and point out that the wage rate sometimes is inadequate.

The starting point in the labor supply literature is sickness absence coupled with labor demand. The few economic models about presenteeism are explored and found lacking in the sense that they do not capture the essence of presenteeism. However, discrete choice models (random utility models) seem to be adequate in that the choice about going sick to work basically is a discrete choice situation that can be extended to include discrete counts, i.e. episodes of presenteeism within a given time period.

The econometrics of presenteeism must have count models as the starting point due to the many zeroes, i.e. many persons do not experience presenteeism and, if they do, usually relatively few days (‘events’) in a given period and the discrete choice nature of presenteeism. Drawing on the econometric literature on utilization of medical services, the following models are discussed briefly: Poisson models, negative binominal, zero-inflated negative binomial, two part models (hurdle models) and latent class models (finite mixture models). This is in contrast to almost all previous literature where logistic regression has been the dominant statistical strategy. The Poissson model is discarded because an important feature (mean – variance) does not hold. The other models are all used in the empirical part of the paper, and an attempt at model selection is made.

The empirical analyses are based on a cross-section survey of Danes in the labor force, N=4,060. The survey was designed with presenteeism in mind – one of the few available data sets at present. Ideally, theory/models should guide empirical work, but can do so only if fully specified theories are available and this is not the case for the random utility models that do no provide much guidance on relevant explanatory variables. The explanatory variables therefore are selected from the existing empirical works along with a number of new variables used in the survey, e.g. attitudinal variables about presenteeism and sickness absence and questions about work environment.

A consistent result across all analyses is – not surprisingly - the importance of self reported health status: The worse health situation, the more presenteeism. . Another consistent result is that sickness absence and presenteeism are positively correlated. Persons with managerial positions also consistently have more presenteeism Age and genders are also (almost) consistently statistically significant. Fear of unemployment is also consistently and significantly related to presenteeism.

JEL Classifications: J22, J24, I12, C35

Keywords: Presenteeism, sickness absence, labor supply, cost-of-illness, economic evaluation, count models

Table of contents

Introduction[a]

Presenteeism describes a situation where employees are on the job but, because of illness, injury, or other health related conditions, they are not functioning at peak levels. In a sense it is the flip side of sickness absenteeism, where employees do not come to work for health related reasons – with presenteeism they are on the job despite some health problems.

Presenteeism attracts increasing attention in the research community and more fleetingly in the policy sphere. The literature is largely empirical without much formal theory or modeling. With the exception of Johns from organizational theory[1], the only attempts to provide a possible theoretical model framework has been provided by economists[2-5 6-9].

From an economic point of view there are two approaches to presenteeism. One is to look at the costs associated with presenteeism, typically lower productivity. This has been explored at some length[3-5, 10-13]. The other approach is to look at presenteeism from the perspective of labor supply and the demand for labor.

There are two aims of the present study, namely to explore possible economic models of presenteeism and to explore the issue econometrically based on a rather unique Danish survey data set. As an aside the determinants of productivity loss are also explored.

The remainder of the paper is organized as follows. The section on background sketches the two dominant definitions of presenteeism and the related research traditions. In the following section possible theoretical frameworks are investigated followed by a description of the data used in the empirical part followed by a small section on econometric methods. This is followed by two sections: One on descriptive statistics and one on estimation results for determinants of productivity changes and presenteeism days. Discussion and perspectives close the paper.

Background

Presenteeism can be defined in two ways, reflecting at the same time two approaches to the study of the phenomenon[b].

On the one hand presenteeism is defined in terms of productivity losses. This is mainly the US-approach. This is a somewhat circular approach in that a phenomenon is defined in terms of one of the important consequences. On the other hand presenteeism is defined in terms of going to work despite being ill/unwell, where the employee might alternatively

[a] Comments from participants at the Workshop on absenteeism, University of Trier, October 2013, are gratefully acknowledged

[b] See Johns table 1 for a list of nine definitions and a discussion of their scientific utility[1].

have called in sick, but for at variety of reasons did not do so. This is the typical (northern) European definition and approach where interest if focused on reasons for and (personal) consequences of presenteeism rather than the productivity losses.

Defining presenteeism by means of consequence for the employer Burton et al. in 1999 described presenteeism as follows – probably one of the first times the term presenteeism was used in a scientific article:

> "However, absenteeism and disability costs should be recognized, at best, as a significant contributor to an incomplete estimate of the total loss of productivity resulting from health impairment. These costs only provide a partial measure of the total lost productivity for a group of employees whose health problems are so severe as to prevent them from working. What are seldom measured are the decrease in productivity for the much larger group of employees whose health problems have not necessarily led to absenteeism and the decrease in productivity for the disabled group before and after the absence period. This decrease may be captured by a measure of "presenteeism," the decrement in performance associated with remaining at work while impaired by health problems. Presenteeism could be measured in costs associated with decreased or slowed output, failure to maintain a production standard, additional training time, errors in work, substandard output, and other events." [14]

The research based on this approach then has mainly been concerned with two related issues: 1. Developing instruments to measure (self reported) productivity as it relates presenteeism, e.g. the Work Limitations Questionnaire and the Stanford Presenteeism Scale[15] as two examples out of at least 14 instruments[16], and 2. the estimation of the costs of diminished productivity, e.g. research by Goetzl et al[17]. Underlying the estimation of the economic loss lays the question of the 'correct' valuation method, e.g. the wage rate[5, 9, 18, 19].

Academically the productivity approach can be tied to the literature on cost-of-illness[20] and economic evaluation[21] and the indirect cost component of such studies[5, 9]. The indirect costs relate to labor market consequences of particular illnesses, for instance sickness absence related to rheumatism or allergy, where the costs of presenteeism can be seen as a new component in addition to absence and disability. With the exception of Brouwer, Koopmanschaf and Pauly[3-5, 8, 9] the economic valuation issue has not been explored theoretically.

Aronsson et al [22, 23] working in the epidemiologic approach to presenteeism describe the emergence and definition of presenteeism in the following way:

> "The changed worklife climate of the 1990s [… in Sweden] has also turned sickness presenteeism into a topical subject. The concept has been used to designate the phenomenon of people, despite complaints and ill health that should prompt rest and absence from work, still turning up at their jobs"

However, presenteeism in the sense defined probably has existed at all times, but research attention has emerged over the past 10-20 years.

The epidemiologic tradition has been concerned with an empirical a-theoretical search for causes and consequences of presenteeism, e.g. personal circumstances, job characteristics and possible later negative health consequences of presenteeism.

There is a scarcity of work on presenteeism in economics. For instance, a search in EconLit using search phrases ‘presenteeism’, ‘presenteeism AND labor supply’ and ‘presenteeism AND absenteeism’ led to 16, 9, and 10 items respectively before excluding largely irrelevant items, thus demonstrating the unexplored nature of presenteeism in (health) economics.

The medical/epidemiological literature is far larger, but far from overwhelming. Searching PubMed with the same phrases as for Econlit obviously gave another result: 346, 3, and 266 respectively, but a perusal of the list shows several articles of marginal interest. This literature of course has a different focus than in the economics literature, including absence of a theoretical framework.

A similar search of Web of Science before sorting and exclusion led to the following numbers 456, 0, and 235 for the same search phrases as for EconLit and PubMed

Johns recent review adds the interest of organizational researchers[1] but the following of the organizational approach at present is limited even though Johns was very clear: “…the presenteeism phenomenon is too interesting and too important for theoretical and practical reasons to be left in the sole hands of medical researchers and health care consultants”, p. 537 [1]

Theoretical framework

As mentioned above from the perspective of economics two different lines of research can be distinguished: 1. The economic loss associated with presenteeism, i.e. lower productivity of employees, embedded in the theory of economic evaluation and the cost-of-illness literature, and 2. Labor supply and demand issues as viewed from the perspective of employees and employer respectively. While absence reduces labor supply, presenteeism on the other hand maintains the level of supply, but possibly at a lower level of effort.

Indirect costs in economic evaluation and cost of illness studies, COI

Much of the interest in workplace productivity stems from the evaluation of pharmaceuticals as illustrated by possibly one of the first studies (1996) taking an interest in work place productivity consequences of migraine[24] and the possible mitigating effect of a migraine drug, e.g. with the new drug lost productivity was 30% lower compared to the control therapy. Among other things, patients recorded time missed from work because of migraine symptoms, time worked with migraine symptoms and percent effectiveness while working with migraine – themes found in all subsequent literature. This in turn was inspired by a 1992-study where patient data on reduced work productivity also was recorded[25]. The study was conducted by researchers from one of the big pharmaceutical companies that for obvious reasons were interested in the plethora of possible effects of new drugs.

The idea of presenteeism was first formulated in terms of economic evaluation in 1999 when Dutch researchers – without using the term presenteeism - noted that in economic evaluations of health care interventions[3], indirect non-medical costs or productivity costs often play an important role if the intervention in question concerns people with paid or unpaid work. The idea of looking at absence and disability was well established in the literature while inclusion of presenteeism costs was largely neglected – in part due to measurement problems. While absenteeism is recorded by the work place, presenteeism – a bit in the nature of things – is not recorded on a regular at the place of work, if ever.

The incorporation of indirect costs (sickness absence, disability and premature death) in economic evaluations has been – and still is - much debated[5, 21, 26]. Apart from Brouwer and Koopmanschaf[3], economists have not been much involved in measurement of presenteeism, but rather taken an interest in the economic valuation of these productivity losses [5, 8, 9, 18, 27]. There has not been much dialogue between economists and medical/health service researchers.

Economic theory of labor supply/demand and presenteeism

Presenteeism so far has only been addressed theoretically in two articles [6, 7]. None of them are or will be definitive pieces of work.

In both articles absenteeism and presenteeism are tied together and naturally coupled to labor-supply, e.g. absenteeism reduces labor supply while presenteeism adds to/maintains labor supply, albeit at lower level of effort. However, the traditional labor supply models are inadequate for a number of reasons. Therefore both articles also include the demand side, i.e. the perspective of place of work (company). The following is an ultra short critical survey starting with models of absence that have been the starting point for models of presenteeism.

Models of sickness absence

In the first review of the economics of absence[28] from 1996 Brown and Sessions noted that the area was underdeveloped relative to other areas of labor economics. They went on to note that in the models of absenteeism based on the traditional static neoclassical labor supply theory (work – leisure choice) absenteeism essentially was based on the premise that it arises not because the individual is unable to work, but because he/she chooses not to, i.e. absence is voluntary and due to an attempt to adjust, if possible, to a utility maximizing position[c] It is a striking weakness as of 1996 that theoretical models of labor supply ignored health status of the individual[d] - and for that matter other determinants of sickness absence. Empirical work by economists is not always based on an explicit theory or, if the case, standard labor supply theory, e.g. Allen's 1981 classic[29]. The model does not include health status.

Over the past 10-15 years much progress has been made in the economics of absence. As is to be expected much of the literature focuses on the effect of economic incentives, i.e. either within an efficiency wage[e] framework or focusing on the payment/remunerations structure and/or degree of compensation in case of sickness absence – and hence within the traditional choice and incentive framework – disregarding for instance accidents at work and the like, i.e. involuntary absence: "The analysis of sickness absence is placed firmly in the agenda of economics by the idea that sickness absence is the consequence of choices that are mediated by economic (and other) incentives"[32].

The theoretical sickness absence/presenteeism models can be grouped into three main (somewhat overlapping) groups[33]: 1 The supply side approach, 2. The efficiency wage approach, and 3. The contract approach. The latest addition is a model type based on the health capital/production of health[34] .

The neo-classical labor supply approach has already been outlined above. The main point is that sickness absence modeled within the work-leisure framework is a choice/adaption variable, in part due to working hours being fixed exogenously, e.g. through union contracts or legislation. If more leisure is desired this is done through sickness absence meaning that absence is shirking and not rooted in underlying health problems or accidents at work. Within this framework the existence of compensation for sickness absence reduces labor supply.

Chatterji and Tilley[7] noted in connection with the classical model that it would be natural to introduce some index of health status, θ, so that lower values of θ were associated with

---

[c] Allen illustrates this clearly: "When a worker contracts for more than his desired hours given w, he retains an incentive to consume more leisure. One way of doing this is to be absent from work." (p. 78)[29]. Economists are amazingly naïve – with greater faith in models than obvious 'real world' observations.
[d] The earliest exception is probably Barmby[30] who in an attempt to move away from the supply-orientation introduced employer monitoring of effort, and hence absenteeism/shirking. To this end he introduced asymmetric information regarding the health status of the employee.

[e] For the sake of clarity, following the New Palgrave Dictionary of Economics, 'efficiency wages' is a term used to express the idea that labor costs can be described in terms of efficiency units of labor rather than in terms of hours worked, and that wages affect the performance of workers. The incentive effects of wages stem from the effect of the level of compensation on the cost to the worker of being fired. Thus, wages above the market clearing level will increase effort, decrease employee theft, decrease absenteeism, and decrease quits. – The classic article is the 1984 shirking model by Shapiro and Stiglitz [31] where the problem is posed in terms of moral hazard. In these models absence is supposed to reveal the employee's level of effort.

healthier states. However, no formal models absenteeism/presenteeism following this idea have been identified[f].

It is commonly assumed[38] that utility and the marginal utility of consumption (income) are decreasing in health status. As health status improves, i.e. θ decreases, the employee places more value on leisure relative to consumption and the indifference curves become steeper. Despite the impeccable model logic this, however, runs counter to intuition. Intuitively one would expect that sickness absence, i.e. true health related absence, would increase with increasing values of θ and decrease with increasing values and hence increasing labor supply. However, this model assumption also rests on somewhat shaky ground[39][g].

The supply approach has been criticized for having passive companies, i.e. demand reactions are absent. While this is true, it is nevertheless the employee who makes the decision to show up for work.

Barmy et al[30] were among the first to recognize that employees may be absent with or without good cause. They used the efficiency wage approach whereby, among other things, the actions of the employer could be modeled. A more recent example of the wage efficiency approach is the work by Ose[40]. In her model she tries to separate the effects of voluntary absence and absence related to ill health, where health effects are assumed to be tied to working conditions At the general level her model builds on and extends the classic 1984 Shapiro and Stiglitz efficiency wage model[31].

The contracting approach goes back to Coles and Treble[41] who looked at the sickness absence from the employer perspective. Workers can be either absent with cause, choose to be absent without cause or choose to be at work. The employer can only observe the absence-attendance choice of the employee. The challenge for the firm is to choose some wage-sick pay contract so as to maximize profit subject to a zero profit condition and an incentive compatibility constraint.

Like in the other two approaches the focus is essentially on economic incentives and asymmetric information. Other causes of absence are not really included, e.g. the working environment (physical and mental).

---

[f] We are disregarding here Gary Becker's allocation of time approach[35] to labor supply and the human capital extension to the production of health[36]. This framework has been used to model 'the demand for absence', and been applied using a Danish data set[37]. Podor and Halliday recently analyzed health status and allocation of time in a framework that explicitly disregarded the Grossman approach. They noted that "There are key differences between this model [… the two authors'] and the canonical model of health investment discussed in Grossman. In that model, time allocation also plays a crucial role. There is a construct called 'sick time' that is essentially a black hole that encroaches upon a person's stock of time that can be allocated to either leisure or production. Importantly, health does not impact productivity", footnote 3.

[g] Finkeltstein et al [39] notes that "If the shape of the utility function varies with health status, this will affect the economic analysis of a number of central problems in public finance, including the optimal structure of health insurance and optimal life-cycle savings. We define health state dependence as the effect of health on the marginal utility of a constant amount of nonmedical consumption. A priori, the sign (let alone the magnitude) of any health state dependence is ambiguous".

Models with presenteeism

In their 2004-article Brown and Sessions extended the model of absenteeism developed by Barmby to include presenteeism. The approach was basically the work-leisure-income model utilizing something akin to a (health) dependent health state[h]. Utility is an increasing function of income and leisure depending on some parameter, θ, representing the general level of health. θ is increasing in sickness and is randomly distributed over the interval [0,1]. Individuals value leisure time more as θ --> 1, essentially meaning it becomes increasing onerous to supply effort at higher levels of sickness.

Chatterji and Tilley[7] develops a model that includes both presenteeism and absenteeism. It is done within the contract framework using principal-agent thinking and looking at full information and perfect contracting and incomplete contracting.

The scenario modeled is a principal (employers) wanting to run a one-time projects by employing agents (employees). Output is assumed to depend on both the attendance and health, θ, of the agent/employee. There are only two possible states of health, ill and healthy, that are exogenously determined. (this may be a reasonable assumption in the short term, but not medium and long term, where health economists would assume that health status largely is an endogenous variable).

It is assumed that the attendance of employees with low health status lowers output "and thus firms may wish unhealthy workers to stay home", p. 676. This is buttressed by a quote from Labor Research (August 1998): "the lowest possible absence rates are not necessarily the best outcome for a firm …, as this may result in presenteeism. This means people coming to work when they should be at home working below par, putting themselves at increased risk, passing on illnesses to their workmates and undermining morale at work". They then develop a model, where employers attempt to separate the workforce according to θ and pay sickness absence compensation to those with absence.

The model totally abstracts from institutional details, e.g. legislated sick pay paid by the employers for at period of time (in Denmark at present the first 5 weeks of a sickness absence period). Usually sick pay initially is paid at the level of the regular pay.

---

[h] Formally we have a state-dependent utility function. The common senses idea is that expected utility from consumption in general is not the same if one is well as when one is ill or that the value of health care is not the same if one is well as when one is ill. "In other words, the utility of consumption (or investment) is dependent on one's own state (of health, amongst other things) and external characteristics (an ice cream in the rain is not the same thing as an ice cream in the sunshine). The idea also introduces another role for uncertainty, viz. uncertainty about one's health state (how high is my 'bad' cholesterol? Am I a carrier of that gene?) or about the weather (etc.). Oddly, the idea is much less used and discussed in health economics than one might expect" [42].

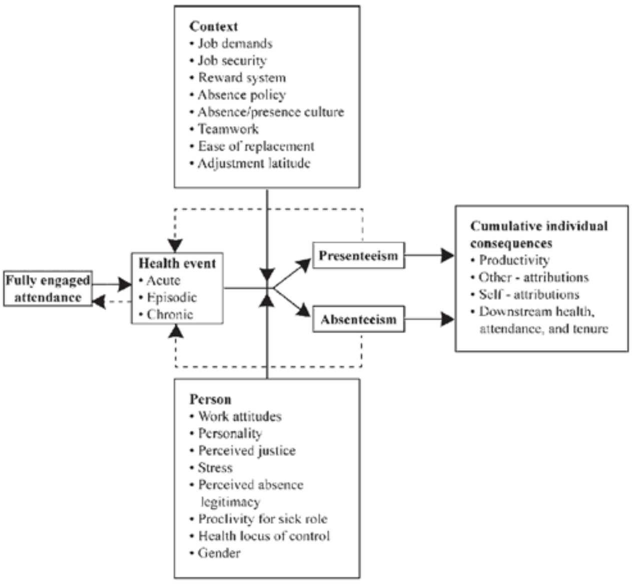
In addition it is also based on the assumption that the decrease in productivity amounts to more than the value of the production lost due to sickness absence. Obviously this is hard to believe unless the underlying reason for presenteeism is a contagious disease.

What should a model of presenteeism contain and ideally provide answers to?

A theory[i] is more than a collection of possible explanatory variables, but this may, however, be a first step towards constructing a theory and start focusing on important phenomena that a (mature) theory should address. Johns in his review of the presenteeism literature presents figure 1

From an economic perspective the question is what affects the choice between presenteeism, sickness absence and allocation of time, including work time, and the possible economic incentives and the role of health. Given the relevant institutional labor market details, i.e. sickness compensation (the same as the at work pay? In Denmark this at present is the case for 5 weeks of continuous sickness absence). This of course makes it harder to identify relevant economic incentives. Hence, a different approach must be taken as discussed in the next section on a discreet choice model

Figure 1: Components of a model of presenteeism (and sickness absence)



Source: Johns, figure 1 [1]

[i] For the sake of clarity: Economists often speak of models. A model basically is a theoretical construct representing economic processes by a set of variables and a set of logical (usually mathematical or geometrical) relationships between them. Figure 1 and 2 are examples of this line of thinking and the two models of presenteeism using demand and supply of labor is based on mathematical modeling. An economic model is a simplified framework – often too simple - designed to illustrate complex processes. – What is the difference then between ‘model’ and ‘theory’? Economists often use the terms as synonyms, i.e. theories as models.

A model based on the random utility model (discrete choice)

In the section below on count models it is noted that they have a behavioral interpretation. This is used in the following to explain initially how a binary decision: go to work (presenteeism) or stay at home, can be understood within a random utility framework (discrete choice models) that maps into a binary statistical model, see Greene[43], chapter 17 for the binary choice model and chapter 18 for multinomial choice, ordered choice and event counts. The basic ideas were developed by nobel prize winner Daniel McFadden. His nobel lecture gives a good overview[44].

Compared to the continuous leisure – work choice models described above, we are concerned with choice among a discrete set of alternatives. Discrete choices then are a contrast to standard choice models in which the quantity of each good or hours worked is assumed to be a continuous variable. Rather: Do I go to work or not – not how many hours do I (want to) work. This seems far better suited to sickness absence and presenteeism[j] - and has been used for modeling labor supply[45, 46].

The essential idea behind random utility models in general is that revealed behavior, i.e. going to work or not, is based on utility maximizing behavior. Consider an individual agent choosing a single option among a finite set of alternatives, in the limiting case only two alternatives: go – or not go to work. This is the realm of behavior that is considered in random utility models, RU or RUM.

In RU models preferences for such discrete alternatives are determined by the realization of latent indices of 'attractiveness', i.e. utility. Utility maximization is assumed to be the objective of the decision process and leads to observed choice in the sense that the agent chooses the alternative for which utility is maximal. Individual preferences depend on characteristics of the alternatives and the tastes of the agent.

An RU model defines a mapping from observed characteristics into preferences. The analyst however cannot observe all the factors affecting preferences and the latter are treated as random variables. By its abstraction from various idiosyncratic factors, the model uses stochastic assumptions to describe unmeasured variation in preferences.

An operational way to allow for maximization of latent preferences is to consider a utility function that is decomposable into two additively separable parts, (1) a deterministic component V specified as a function of measured attributes of the alternatives and/or the individual, and (2) a stochastic component ε representing unobserved attributes affecting choice, interindividual differences in utilities depending upon the heterogeneity in tastes, measurement errors, and functional misspecification

[j] Technically it should be noted that in the continuous case, calculus methods (e.g. first-order conditions) can be used to determine the optimum amount chosen. Discrete choice analysis examines situations in which the potential outcomes are discrete, such that the optimum is not characterized by standard first-order conditions.

Turning this into a formal model: let

(1) $U_{ij} = V_{ij} + \varepsilon_{ij}$.

be the utility of alternative j for agent i, where V_{ij} is the deterministic component and ε_{ij} the random component.

The deterministic component V_{ij} most often is assumed to have an additively separable linear form – fitting nicely into a regression framework: $V_{ij} = \mathbf{x}_{ij}'\boldsymbol{\beta}$ where $\mathbf{x}_{ij}$ and $\boldsymbol{\beta}$ are the vectors of exogenous variables characterizing the agent and the alternatives and parameters, respectively.

In the hypothetical case that V contains perfect information about the determinants of utility, the consumer would simply choose the product with the highest V_{ij} The stochastic terms ε_{ij} shaping the true and latent utility in (1), introduce uncertainty regarding the choice and therefore, choice probabilities are invoked to describe choice behavior.

The probabilistic description of choice has been introduced not to reflect that behavior is probabilistic. Rather, it is the lack of information that leads to treating utility as a random variable and therefore describes choice in a probabilistic fashion. In fact, the properties of RU models can be attributed to the specific assumptions that each model implies about the stochastic terms.

Under the utility maximization rule, a consumer facing a set of available alternatives C = {1,2,3, . . . ,M} will choose an alternative j with probability $P(j) = P(U_j > U_k)$ for all $k \in C$, $k \neq j$,
or as it follows from (1):

(1) $P_j .= P(\varepsilon_k < V_j - V_k + \varepsilon_j)$. for all $k \in C$, $k \neq j$:

The probability that j is chosen is then obtained by making assumptions about the form of the distribution of the random variables and integrating Eq. (2) over a continuum of all possible values for ε_j.

Now, turning to presenteeism and two alternatives. Denote by U_1 the utility of going to work despite feeling unwell, and by U_0 the utility of staying at home. Both U_1 and U_0 are latent variables for person i modeled as follows

(3) $U_{1i} = \mathbf{x_i}'\boldsymbol{\beta}_\mathbf{i} + \varepsilon_{1i}$
(4) $U_{0i} = \mathbf{x_i}\ '\boldsymbol{\beta}_\mathbf{0} + \varepsilon_{0i}$

where $\mathbf{x_i}$ is a vector of individual attributes (and one can easily just add another vector for the attributes of the alternative – but omitted here), and ε_{0i} and ε_{1i} are random errors.

Then for person i who goes to work (presenteeism), we have by implication:

(5) $U_{1i} > U_{0i} \rightarrow \varepsilon_{0i} - \varepsilon_{1i} < \mathbf{x_i}'(\boldsymbol{\beta_1} - \boldsymbol{\beta_0})$

For individual i who goes to work $U_{1i} > U_{0i}$ we then observe a value of 1 on the observed outcome-variable y_i =1, i.e. presenteeism, otherwise we observe $y_i = 0$. The probability that $y_i = 1$, is given by $P[\varepsilon_{0i} - \varepsilon_{1i} < \mathbf{x_i}'(\boldsymbol{\beta_1} - \boldsymbol{\beta_0})$. In other words, the probability of the decision to go to work is characterized by a binary outcome model.

In other words, the observed choice between the two alternatives reveals which one provides the greater utility, but not the unobservable utilities. Hence, the observed indicator, called y above, equals 1 if $U_1 > U_0$ and 0 if $U_1 \leq U_0$.

This line of thinking covers binary choices, multinomial choices, ordered choices, and event counts, where the observed outcome is a count of the number of occurrences, e.g. days or episodes of presenteeism/sickness absence. In many cases, this is similar to the preceding three settings in that the "dependent variable" measures an individual choice, such as the number of visits to the physician or days of absence or presenteeism. This obviously is of particular interest here, see Greene chapter 18, section 18.4[43] using the Poisson[47], negative binomial and the two part (hurdle) regression model.

It should be noted that only *utility differences* matter. The probability that person i chooses a particular alternative is determined by comparing the utility of choosing that alternative to the utility of choosing other alternatives:

$P_{ij} = P[y_{ij} = 1]$

$= P[U_{1i} > U_{0i}]$

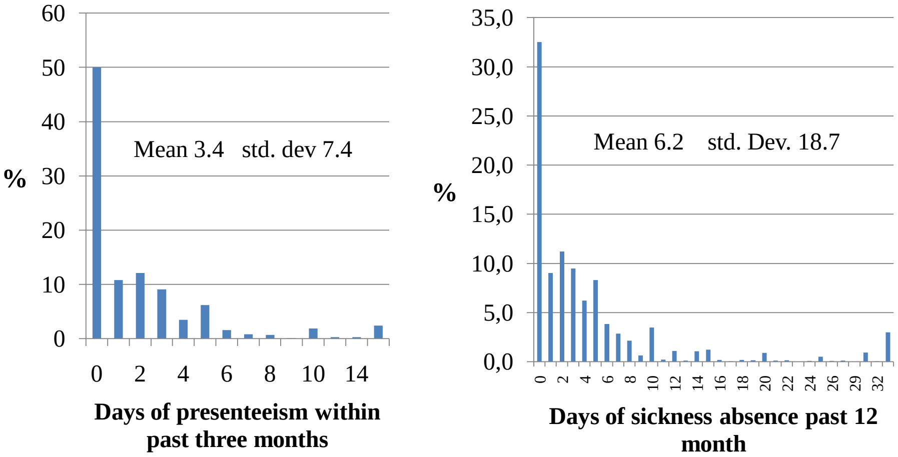
$= P[U_{1i} - U_{0i} >]$

The last term shows that the choice probability depends only on the difference in utilities between alternatives, not on the absolute level of utilities. Equivalently, adding a constant to the utilities of all the alternatives does not change the choice probabilities.

Econometric approaches

The majority of analyses of determinants of presenteeism have used logistic regression to analyse the relationship between presenteeism, however defined, and independent/explanatory variables, e.g.[22, 23, 48]. However, this entails not only loss of information, but also a loss of a better understanding of presenteeism.

Figur 2: Illustration of dominance of zeroes in data on sickness absence and presenteeism (the Danish presenteeism Survey).



Count-models

Statistically a dominant feature of presenteeism and sickness absence data is many zeros, i.e. many persons do not experience presenteeism or sickness absence and a discrete number of counts, figure 2. The question is what the relevant econometric model is in such a case. In health economics this issue is also encountered when analyzing utilization data, i.e. hospitalizations or doctor visits. Quite a bit of work has been carried on the econometrics issues, [49-53] whereas the issue is at the early stages of exploration for presenteeism data[54, 55]. The following models have been discussed: Poisson models, negative binominal, zero-inflated negative binomial, two part models (hurdle models) and latent class models (finite mixture models), commonly named count models[56].

The Poisson model which handles the probability that a given number of (rare) events like sickness absence or presenteeism occurs during a time period. But the Poisson distribution does not really suit presenteeism/absence data: it assumes equality between mean and variance, whereas presenteeism/absence data are very often over-dispersed, meaning characterized by a variance that is significantly higher than the mean.

When very few employees show very long absences/presenteeism, the mean still remains low as the variance is high. That could happen, for instance, if the population bears heterogeneous health. This effect could be taken into account by a negative binomial model that actually corresponds to a Poisson model with unobserved population features (health, for example). Contrary to the Poisson model, it does not imply equality between the mean and the variance.

The zero-inflated negative binomial, the two part model and finite mixture model all assume two separate statistical processes (and obviously also relevant behavioral explanations of the two processes, e.g. discrete choice as outlined above). One for the zero group and one for the non-zero group.

Zero-inflated negative binomial regression, ZINB, is for modeling count variables with excessive zeros and it is usually overdispersed, i.e. variance is bigger than the mean, count outcome variables. Furthermore, the excess zeros are generated by a separate process from the count values and that the excess zeros can be modeled independently[k].

In the health econometrics of the demand for health care it is common to use the two-part model (the hurdle model), TPM. The first part of the TPM is a binary outcome model, typically a logistic model that describes the distinction between persons without and with presenteeism or sickness absence. The second part describes the distribution of non-zero presenteeism conditional on some presenteeism modeled as an integer-valued random variable. The second part typically relies on the negative binomial model, and the term hurdle model is used for this two-part model. However, other possibilities than the negative binomial model, e.g. the standard Gaussian model, exist.

In contrast to ZINB model it is recommended that the explanatory variables in the two parts are the same. Frequently the first part of the ZINB-model only includes a subset of the independent variables in the second part. The reason for this, however, is unclear.

The sharp dichotomy between users and non-users in utilization studies (or attenders and non-attenders if we focus on sickness absence and presenteeism) in the TPM has been challenged by among others Deb & Trivedi, and d'Uva[49-53, 58, 59].

---

[k] On www.statisticalhorizons.com/zero-inflated-models there is an illuminating discussion between the two well-known econometricians/statisticians William Greene and Paul Allison. The latter is skeptical of ZINB whereas the first named has been a leading proponent. Allison claims that " It's certainly possible that a ZINB model could fit better than a conventional negative binomial model regression model. But the latter is a special case of the former, so it's easy to do a likelihood ratio test to compare them (by taking twice the positive difference in the log-likelihoods). In my experience, the difference in fit is usually trivial." Greene notes in response to Allison's argument that that "The zero inflation model is a latent class model. It is proposed in a specific situation – when there are two kinds of zeros in the observed data. It is a two part model that has a specific behavioral interpretation (that is not particularly complicated, by the way)." - Without entering into the debate it should noted that it is of course possible to rely on statistical tests to choose between two competing models. But more importantly is to what extent the models mirror a theoretical understanding of the phenomenon at hand. This coincides with the following observation from public health: "In general, for public health studies, we may conceptualize zero-inflated models as allowing zeroes to arise from at-risk and not-at-risk populations. In contrast, hurdle models may be conceptualized as having zeroes only from an at-risk population. Our results illustrate, for our data, that the ZINB and NBH models are preferred but these models are indistinguishable with respect to fit. Choosing between the zero-inflated and hurdle modeling framework, assuming Poisson and NB models are inadequate because of excess zeroes, should generally be based on the study design and purpose" [57].

They claim that a more tenable distinction for typical cross-sectional data may be between an “infrequent user” and a “frequent user” of medical care – or in the present case short and long term sickness absence - the difference being determined by health status, attitudes to health risk, choice of life-style etc. depending on the type of data being analyzed. In their proposed alternative model, the latent class model, LCM, there is no distinction between users and non-users of care/attenders and non-attenders. Instead there is a distinction between groups with high average demand/absence/presenteeism and low average demand/absence/presenteeism based on two (or more) latent classes.

Deb and Trivedi hypothesized that the underlying unobserved heterogeneity which splits the population into latent classes is based on an individual’s latent long-term health status which also happens to be relevant for presenteeism. Proxy variables such as self-perceived health status and chronic health conditions may not fully capture population heterogeneity from this source. Consequently, in the case of two latent subpopulations, a distinction may be made between the “healthy” and the “ill” groups, whose demands for medical care (or sickness absence/presenteeism) are characterized by low mean and high mean, respectively.

The mixture/latent class approach then can be interpreted as allowing for latent groups/classes in the population. The data for each group may be characterized by a parameter vector – the same for both groups. Since the group to which an individual belongs is not observed directly, a mixing probability is used to classify individuals probabilistically. The mixture negative binomial model has the virtue of being conceptually simple In the following estimation two latent groups are assumed (‘short’ and ‘long’ term sickness/presenteeism).

In sum then: For the analysis presenteeism and sickness absence data there is a choice between the following models: negative binominal regression, zero-inflated negative binomial regression, two part models (hurdle models) and latent class models (finite mixture models). The choice must be a combination of theoretical considerations about the phenomenon (presenteeism and sickness absence) and statistical considerations, cf. footnote 7. However, as noted in the section on theoretical economic models, coherent and reasonably encompassing models of presenteeism/sickness absence are lacking both in economics and public health.

Data

Data for the following analyses was collected through an internet-based survey aimed specifically at presenteeism (and employee paid health insurance). The presenteeism survey (PRS*)*, is a cross sectional survey of the occupationally active Danish population. It was carried out in December 2010. The realized sample size was 4,060. It is one of the few – maybe the only one – where the main purpose of data collection was to look into presenteeism. Most previous empirical work has been done on data sets collected for other purposes.

Respondents answered a questionnaire aimed at presenteeism ('sick at work') and absenteeism. There was a host of questions about type of work place, type of work, attitudes towards presenteeism/absenteeism, health status along with the usual socio-demographic variables

The data appears to be reasonable representative of the occupationally active Danish population. Approximately 86% of the adult Danish population has internet access at home. The remaining 14% largely belong to the elderly/old part of the population. Since the sample only included occupationally active adults, largely 65 years of age or less, the use of the internet does threaten representativity.

As regards representativity we compared the respondents with the employed background population in Denmark using data from Statistics Denmark. There was a slight overweight of respondents from the capital region (about 2.5%) However, when looking at age and sex males in the age bracket 18-25 were strongly underrepresented (7 percentage point) compared with the employed background population. The same was found in the female age group, but not quite as outspoken. Apart from this it seems that the study population is representative of the Danish working population,

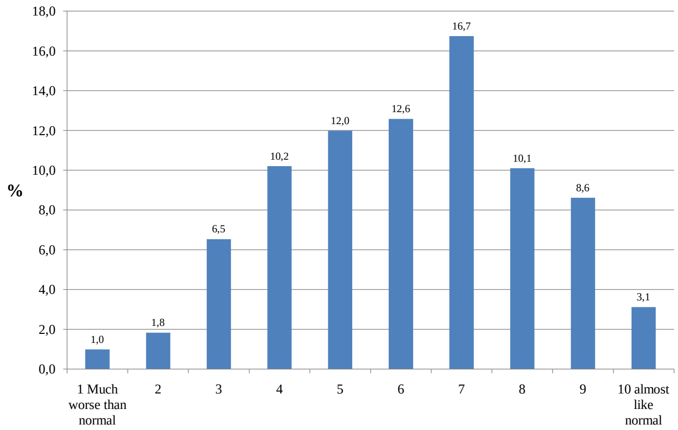
Empirical results for productivity

In the Danish survey the questions about work performance during presenteeism were created by modifying questions from the Stanford Presenteeism Scale, SPS, the WHO questionnaire on work performance[15, 60-62]. While the main focus here is presenteeism days the issue of productivity has also been discussed above and hence is also addressed empirically here. However, not at the same depth as presenteeism days.

Figure 3 shows the response to a partially self-constructed question about overall subjective evaluation of work performance during presenteeism days over the past three months where respondents rated their performance against normal performance on a scale from 1 (much worse than normal) to 10 (almost normal level).). The method, however, is quite similar to the Quality and Quantity method (QQ) which have been used by Brouwer et. al.[3]. The method measures the quantity of production on a 10-point numeric rating scale where 0 represents "nothing" and 10 represents "normal quantity". The study by Meerding et al it shows that the self-reported productivity in the QQ was significantly correlated, with objective work output (r=0.48)[63].

.

Figure 3: Overall rating of work performance during presenteeism, N=2,018



About 1/3 rated their performance from much worse (score 1) to 5.

Table 1 goes behind the summary answer reported in figure 3. It is noteworthy that while 64% reported that they worked slower than normally only 17% had to postpone tasks. A cross tabulation (not reported here) shows that the majority of those who postponed task also reported working slower – attesting to the consistency of the answers.

Table 1 : "Think of those days during the past three month where you went to work despite feeling ill. Which of the following statements describe your situation during those days"
N=2,108 (those who reported presenteeism), % who agreed with the statement

	%
I had problems concentrating	46
I worked slower than normally	64
I had problems making decisions	13
I had to postpone tasks	17
I had to let others take over tasks	5
None of these	15

Table 2 shows a simple OLS regression of possible determinants of impaired work performance. Ordered logistic regression analysis resulted in roughly the same pattern of results as the OLS-regression. The determinants were selected based on the very sparse literature, common sense and the reasoning behind the questions included in the survey answers about company policies, working conditions etc. played an important role.

It is interesting to note that presenteeism did not very much lead to either postponing task or letting others take over. Also, 'only' about 1/3 of those experiencing presenteeism on a scale from 1 (much worse performance) to 10 (almost like normal) were located between 1 and 5.

Table 2: Possible determinants of work performance for employees with presenteeism within the past 3 months, N=1738, OLS, R^2=0.12. Bold results p>0.05.

Dependent variable : work performance, 1= much worse than normal … 10=almost like usual	Coef.	Robust Std. Err.	t	P>t
Health condition				
No of chronic illnesses ((max of 14)	-.016213	.0470228	-0.34	0.730
Self rated health (health status very good=0)				
Good	-.255474	.1609989	-1.59	0.113
Acceptable	**-.6325326**	**.1866861**	**-3.39**	**0.001**
Bad & really bad	**-.9854874**	**.2972103**	**-3.32**	**0.001**
Socio-demographics				
age	**.0226306**	**.0059282**	**3.82**	**0.000**
Gender (male=0)	.1880529	.1091675	1.72	0.085
Occupation (self employed =0)				
* Skilled	.0357077	.3493064	0.10	0.919
* Unskilled	-.0999029	.3848986	-0.26	0.795
* 'white collar'	.0636726	.2996341	0.21	0.832
* Mix	.4072121	.3759231	1.08	0.279
The work place				
No mgt. Responsibility (0=mgt responsibility)	**-.3018599**	**.1210408**	**-2.49**	**0.013**
Type of company (private=0)				
* Public	.0347036	.1143649	0.30	0.762
Seniority, no. Years	.0107963	.0071091	1.52	0.129
No employees,(1-4=0)				
* 5-9 employees	.2598529	.3061439	0.85	0.396
* 10-19 employees	.1572345	.2994322	0.53	0.600
* 20-49 employees	.3241125	.2827616	1.15	0.252
* 50-99 employees	.1139719	.2899668	0.39	0.694
* 100-249 employess	.4154067	.2923703	1.42	0.156
* 250-499 employees	.0642598	.3101757	0.21	0.836
* > 499 employees	.1960375	.2822489	0.69	0.487
Sickness absence interview	**.4719328**	**.157417**	**3.00**	**0.003**
Satisfaction with place of work (very =0)				
*** rather satisfied**	**-.3358958**	**.1356058**	**-2.48**	**0.013**
*** satisfied**	**-.6024693**	**.1582796**	**-3.81**	**0.000**
*** dissatisfied**	**-.728131**	**.2542336**	**-2.86**	**0.004**
*** rather and very dissatisfied**	**-1.034.041**	**.3518595**	**-2.94**	**0.003**
Type of work/task environment ('always or often'=0, 'occassionally or never')				
* gets help and support from colleagues	.0555229	.1126388	0.49	0.622
* uneven workload so task pile up	.102462	.1201933	0.85	0.394
*** do not get all tasks done**	**.507201**	**.1329565**	**3.81**	**0.000**
* other handle my tasks if absent	-.156371	.1324854	-1.18	0.238

* plan your own work	-.0963037	.1430651	-0.67	0.501
* high work pace	-.0328189	.1252186	-0.26	0.793
* decide my own work pace	-.040312	.1163804	-0.35	0.729
*** important with high work space**	**-.2523507**	**.1239973**	**-2.04**	**0.042**
*** mentally exhausting work**	**.3195891**	**.1424323**	**2.24**	**0.025**
*** physically exhausting work**	**.3067469**	**.1527207**	**2.01**	**0.045**
Constant	5.238.995	.4297467	12.19	0.000

Some noteworthy results are: Not surprisingly, work performance declines with worsening health status. Work performance, given presenteeism increases towards 'normal' with increasing. As regards elements related to the place of work: Having a managerial position compared to non-managerial positions is associated with lower decrease in work performance. Employees who have had a sickness absence interview performed better during presenteeism than those who had not such a interview. Work performance decreased with the degree of satisfaction with the place of work.

Econometric explorations of determinants of presenteeism using count models

Referring to the section on econometrics the following is a presentation of results based on the proposed statistical models. This is followed by a section on model selection following the approach of Deb and Trivedi[51, 56].

The following blocks of explanatory variables are used, cf. figure 1 and the table below.

1. *Health status*: Self-rated health status, longterm conditions, chronic illnesses.
2. *Socio-demographics (age, sex, occupation)*
3. *Attitudinal variables*: important to come to work, attitude towards presenteeism and sickness absence etc.
4. *Type of job and characteristics of the job* (managerial position, commitment to job, team-related, possibility to plan one's own work, helping hand from colleagues if absent, physically or mentally demanding
5. *Type of company and company policies*: public-private, number of employees, absence policy and sickness absence interviews, company attitude towards presence/presenteeism, economic consequences of absence/presenteeism

A recent article[54] developed the following 8 hypotheses where it is striking that health status is not included:

Hypothesis 1: When cost of absence rises, absenteeism decreases and presenteeism increases. *Comment*: In the Danish labor market sickness absence is compensated fully at the going wage/salary for the first three weeks as of 2010 (legislation). To a considerable extent then the hypotheses is largely irrelevant for short term illness. However, the type of remuneration may play a role in that the base for calculating sickness absence may incur an economic loss, see table G and H below.

Hypothesis 2: Absenteeism (resp. presenteeism) is lower (resp. higher) for employees with 'team responsibility'. *Comment*: We have some information on team-work, i.e. colleagues suffer/have trouble if I am absent from work. Hence enabling us to test this hypothesis.

Hypothesis 3: As hierarchical level rises, absenteeism decreases and presenteeism increases. *Comment*: We have an item addressing managerial responsibility and the number of subordinates.

Hypothesis 4: Job insecurity associated with fixed-term contract deters absenteeism and reinforces presenteeism. *Comment:* We have an item related to fear of unemployment.

Hypothesis 5: Both absenteeism and presenteeism could be affected by a job mobility. *Comment*: We have no relevant information on this point

Hypothesis 6: Both absence and presenteeism are (slightly) correlated with children at home. *Comment*: We have information on family situation, including children living at home

Hypothesis 7: Both absenteeism and presenteeism are gender dependent. *Comment*: We have information on gender.

Hypothesis 8: Both absence and presenteeism are age dependent. *Comment*: We have information on age of the respondents

The table below shows expected signs of the regression coefficients with small supporting arguments.

In several cases there can be arguments for both expecting a positive and a negative sign.

It obvious that health status is an important variable, otherwise the whole idea of presenteeism would not make sense. But what should one expect empirically?

For instance, does bad health status lead to less presenteeism compared to those with excellent health status? One might actually argue that those with low self-rated health status would have more presenteeism because they also might have higher sickness absence than those with excellent health and therefore they try to compensate for more sickness absence by demonstrating presenteeism. In the table this would lead to a positive sign. On the other hand there also might be arguments claiming that those with low self-rated health status will have less presenteeism because the find it difficult to cope. – Similarly, if one looks at chronic illnesses. We have asked about the presence of 14 conditions, i.e. allergy, diabetes, migraine, depression, and other psychiatric conditions. But do persons with one or more chronic conditions have more presenteeism than those without? Yes, if one assumes that persons gradually learn to cope with the chronic condtions – but not perfectly. No, if one instead expect that they call in sick, if the chronic condition flares up.

Table 3: **Summary of expected sign of regression coefficient with number of presenteeism days as dependent variable. Zero = reference category for dummy-variables shown. In several cases there may arguments both for plus and minus signs. However, in the fully specified two part models the signs mainly refer to the second part (positive number of presenteeism days)**

Health Status	SIGN	
No of chronic illnesses ((max of 14)	-	Harder to work with more chronic conditions
Self rated health (excellent=0)		
* Good	+	Expect less presenteeism with lower health status
* Acceptable	+	" "
* Bad & really bad	+	" "
Socio-demographics		
Age	+	Have learned to cope with presenteeism with age
Age squared	-	With a declining rate
Gender (male=0)		
Occupation (self employed =0)		
* Skilled		
* Unskilled		
* 'white collar'		
* Not classified		
Attitude to absence/presenteeism		
Ought to stay at home if below normal performance (yes=0)	+	Those who disagree more frequently have presenteeism days compared to those who agree
Wrong to go to work if I infect colleagues, customers (yes=0)	+	" "
By calling in sick you recuperate faster (yes==)	+	" "
Fear of unemployment (to a very high degree =0)		
* high degree	-	Compared to those who have a high degree of fear those with less fear have fewer presenteeism days.
* some degree	-	" "
* lesser degree	-	" "
* not at all	-	"..............................."
The work place	-	"............................... "
No mgt. Responsibility (0=mgt responsibility)	-	Persons with management responsibility have more presenteeism days than those without
Commitment and involvement of work (very high=0)		
* high	-	Compared to those with high commitment those with less will have fewer presenteeism days
* to some degree	-	
* lesser degree	-	
* not at all	-	

Type of company (private=0)	-	
* Public	-	Unclear sign – but expect fewer presenteeism days in the public sector compared with the private
Seniority, no. years	-	Increasing seniority and decreasing presenteeism – gets to know the work place
No employees,(1-20=0)		Unclear sign – but may expect clearer norms and social control at small companies, meaning that presenteeism declines with increasing company size
* 20-99 employees	-	” ……………………………………”
* 100-499 employees	-	”……………………………………”
* > 49920-49 employees	-	”……………………………………”
Sickness absence interview (yes=0)		Compared to those having experienced sickness interviews those with will have more presenteeism
Satisfaction with place of work (very =0)		May be similar to work commitment, hence the more satisfied the more presenteeism
* rather satisfied	-	” ”
* satisfied	-	”………………………………”
* dissatisfied	-	”………………………………”
* rather and very dissatisfied	-	”………………………………”
Type of work/task environment ('always or often'=0')		
* other handle my tasks if absent	-	Lower presenteeism than if work just piles
* plan your own work	-	Lower presenteeism – can catch up later
* physically exhausing work	-	May expect lower presenteeism due less stamina
Sickness absence, days past 12 months	-	Would expect a negative correlation, i.e. more absence less presenteeism

Descriptive results

In addition to the standard descriptive statistics, table 4, for the count regression models to follow, tables A - H below provide detailed insight into the data – all too rare made available to the interested reader to give a better ‘feel’ of the underlying data structure..

The tables show one-by-one some independent variables, e.g. health status, and how the dependent variable (presenteeism days the past three months) varies across the categories of the chosen independent variable, including adjusting for possible gender and age differences. STATA’s adjust routine has been used for the adjustment. Across the tables it is interesting to note how little this adjustment changes the picture.

Table: A: Health status and presenteeism days past three months

Self rated health	Presenteeism days	gender-and age adjusted	N
* Excellent	1.40	1.36	666
* Good	1.90	1.90	2,265
* Acceptable	3.08	3.10	939
* Bad & really bad	6.17	6.20	180

Tables A to D show the relationship between various health variables and presenteeism. A very clear picture emerges: The worse the health condition the more presenteeism days.

Table C: **Arthritis and presenteeism days past three months**

Chronic disease	Presenteeism days	gender and age adjustred	N
not osteoarthritis	2,13	2,13	3,634
osteoarthritis	3,57	3,62	416

Table C: Migraine and presenteeism days past three months

Chronic disease	Presenteeism days	gender and age adjustred	N
ikke migræne	2.18	2.18	3,654
har migræne	3.24	3.18	396

Table D: Number of chronic diseases (max of 14) and presenteeism days past three months

No of chronic conditions, max 14	Presenteeism days	gender and age adjustred
0	1.94	1.84
1	2.14	2.37
2	2.97	2.90
3	3.33	3.43
4	3.92	3.96
5	4.13	4.49
6	8.30	5.02
7	6.25	5.56
8	8.0	6.09

It is often assumed implicitly that that if presenteeism is high, then absenteeism is low, i.e. a substitution relationship. Tables E and figure 4 points towards a different picture, that of complementarity. There is a rather clear positive relationship. It is also clear that absence increases with worse health states dispelling the notion that absence totally is a choice variable.

Table E: self rated health presenteeism and absence days (past 12 months)

Self rated health	Presenteeism days	gender-and age adjusted	Absence days	gender and age adjustred	N
virkelig god	1.40	1.36	2,71	1,10	666
god	1.90	1.90	4,20	4,83	2,265
nogenlunde	3.08	3.10	8,12	8,55	939
dårlig - meget dårlig	6.17	6.20	16,47	12,26	180

Figure 4: Scatterplot of presenteeism days (past 3 months) and absence days (past 12 month)

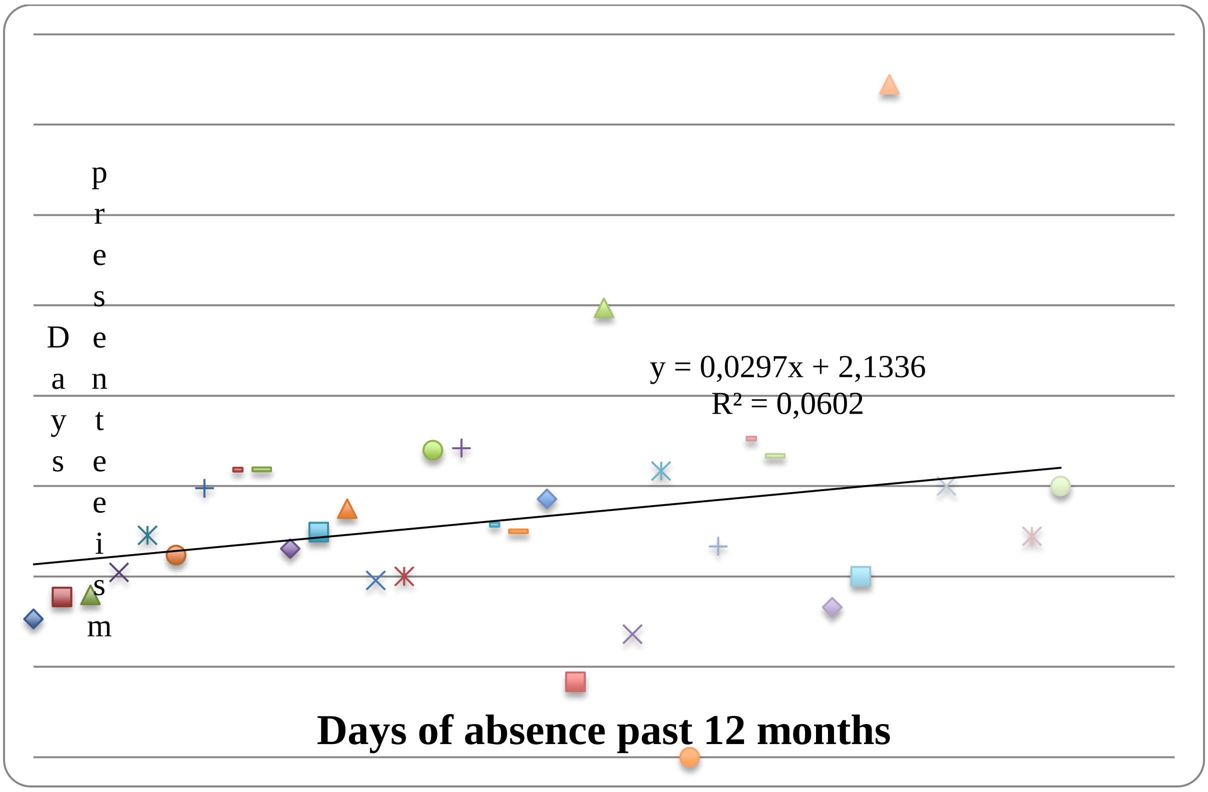


Table F and G shows the results for two of the many work place related data. Having managerial responsibility leads to a greater number of days with presenteeism. So does having had a sickness absence interview (an interview where the employee discusses reasons and possible solutions to a somewhat high level of absence. In Denmark increasingly absence interviews are mandated in union contracts).

Table F: Presenteeism and managerial responsibility

Managerial responsibility	Presenteeism days	gender and age adjustred	N
yes	2.70	2.77	1,128
no	2.12	2.09	2,922

Table G: Presenteeism and absence interview

sickness absence interview	Presenteeism days	gender and age adjustred	N
yes	3.09	3.05	487
no	2.17	2.18	3,563

Due to Danish labor market legislation employees do not suffer economically if they are absent from work (for the first three weeks at the time of the survey. Today it has been changed to 5 weeks). Therefore one would not expect economic incentives to be at play to a significant degree in explaining presenteeism. Hence one would not expect a clear pattern to emerge when looking at presenteeism days and personal pre-tax income, table H. However, on the other hand there may be an influence from the type of remuneration system used at the place and of work and presenteeism, table G. Some union contracts stipulate a maximum hourly pay to be compensated in case of sickness absence.

Table G. Presenteeism and type of remuneration

Type of remuneration	Presenteeism days	gender and age adjustred	N
fixed, typical fixed monthly	2.26	2.20	3,134
hourly pay	1.92	2.39	591
fixed base +piece pay	6.03	2.58	29
fixed pay + commission	1.77	2.77	102
Mixture of the above	3.39	2.96	194

Table H. Presenteeism and personal pre-tax income

Personal pre-tax income	Presenteeism days	gender and age adjustred	N
less100,000 kr.	1.43	2.21	135
100,000 - 199,999 kr,	2.04	2.22	250
200,000 - 299,999 kr,	2.53	2.24	575
300,000 - 399,999 kr,	2.47	2.25	1,105
400,000 - 499,999 kr,	2.06	2.27	723
500,000 - 599,999 kr,	2.10	2.28	319
600,000 - 699,999 kr,	2.32	2.30	207
700,000 - 799,999 kr,	1.50	2.31	95
800,000 - 899,999 kr,	1.89	2.33	73
900,000 - 999,999 kr,	5.06	2.35	35
1000,000 kr.or more	2.03	2.36	64
Do not want to reveal	2.33	2.38	402
do not know	2.5	2.39	76
1 $ = 5.50 kr; 1 euro = 7.50 kr. /January 1, 2014			

Attitudes towards work may influence sickness/presenteeism behavior. Table I shows one of several variables. It is seen that employees making a point of honor always to come to work have a higheder number of presenteeism days compared to those who totally or partially disagree with this statement.

The team hypothesis (no. 2 above) only receives moderate support, table J

Table I: Attitude towards absence and days of presenteeism

Point of honor always to come to work	Presenteeism days	gender and age adjustred	N
totally/partly agree	2,78	2,77	1,719
neutral	2,17	2,24	933
totally/partly disagree	1,74		

Table J: Effect of absence on colleagues and days of presenteeim

Sickness absence affects my colleagues	Presenteeism days	gender and age adjustred	N
totally/partly agree	2.38	2.33	2,648
neutral	1.99	2.23	881
totally/partly disagree	2.28	2.13	521

Table 4: Descriptive statistics for variables in econometric analyses, N=4050

Variable	Average or %	Std err
Dep. Variable: No. presenteeism days	2.27	.0895471
Health Status		
No of chronic illnesses ((max of 14)	.83	.0184308
Self rated health (excellent=0)		
* Excellent	16.4	
* Good	55.9	
* Acceptable	23.2	
* Bad & really bad	4.4	
Socio-demographics		
Age	42.26	.1800203
Age squared		
Gender (male=0)	47.6	
Occupation (self employed =0)		
* Self employed	6.4	
* Skilled	6.9	
* Unskilled	7.7	
* 'white collar'	72.0	
* Not classified	6.1	
Attitude to absence/presenteeism		
Ought to stay at home if below normal performance (yes=0)		
* Yes, agrees	26.5	
* neither/nor - neutral	26.9	
* No, disagrees	36.6	
Wrong to go to work if I infect colleagues, customers (yes=0)		
* Yes, agrees	76.4	
* neither/nor - neutral	15.8	
* No, disagrees	7.8	
By calling in sick you recuperate faster (yes=0)		
* Yes, agrees	59.3	
* neither/nor - neutral	26.9	
* No, disagrees	13.9	
Fear of unemployment (to a very high degree =0)		
* very high degree	6.4	
* high degree	7.5	
* some degree	20.2	
* lesser degree	38.4	
* not at all	27.6	
The work place		
No mgt. Responsibility (0=mgt responsibility)	72.2	

Commitment and involvement of work (very high=0)		
* very high degree	6.5	
* high	14.7	
* to some degree	32.3	
* lesser degree	32.1	
* not at all	14.4	
Type of company (private=0)		
* Public	39.4	
Seniority, no. years	8.0	.1444766
No employees,(1-20=0)		
* 1-19	26.5	
* 20-99 employees	28.3	
* 100-499 employees	20.0	
* > 499 employees	23.6	
Sickness absence interview (yes=0)		
Satisfaction with place of work (very =0)		
* very satisfied	27.0	
* rather satisfied	38.0	
* satisfied	25.6	
* dissatisfied	6.0	
* rather and very dissatisfied	3.5	
Type of work/task environment ('always or often'=0')		
* other handle my tasks if absent	75.4	
* plan your own work	25.8	
* physically exhausing work	78.6	
Sickness absence, days past 12 months	5.5	.1930189

Econometric results for days of presenteeism

Tables 5 to 9 present the results from: Negative binominal regression, zero inflated negative binominal, the two part model and the latent variable (finite mixture models). For the latter two models part one and two are presented in the same table for easy comparison. However, the z-values have been deleted. In all tables results significant at the 5% level or better have highlighted with bold letters.

The tables are organized in terms of the five blocks of variables mentioned above, (p. 21).

Table 10 summarizes the results across the four regression models in terms of sign of the regression coefficients and the significance (≤ 5%).

As would be expected, the two 2-equation models (hurdle/the two part model and finite mixture/the latent variable model) deviate from the single equation regression (to which ZINB a little bit incorrectly is classified).

In the two-part-models there is a clear distinction between non-presenteeism and positive presenteeism pointing towards two separate processes. In some cases signs are reversed between the two equations and the pattern of significance is also changing across the two parts. However, comparing the hurdle model/two part model and the finite mixture model there are also differences, e.g. for occupation where the two part model has significant coefficients in part 2, while the most significant coefficients appear for component 1 in the finite mixture model (and: occupation is hardly ever significant in the single equation models). The same observations holds for commitment to work. This obviously points toward the importance of model selection, i.e. which is the better: the hurdle model or the finite mixture model in that they apparently capture different things in the two parts of the respective models.

Across all the regression models *the health variables*

- all have positive signs (number of chronic diseases) meaning that presenteeism seems to increase with more chronic diseases and that those with self reported health status lower than excellent have more presenteeism than those with excellent health. In addition the coefficients for self reported health status usually are significant.

- The *age and gender* variables also have the expected signs: presenteeism increase with age but at a decreasing rate and women, maybe surprisingly, have lower presenteeism than do males.

- *Fear of unemployment* also has very clear results: those who have little or no fear have lower levels of presenteeism than those with a degree of fear. The results are usually significant with the exception of the second-component in the latent variable model.

- Another consistent result relates to having *management responsibility*. Those with management responsibility consistently have lower presenteeism than those without. Most often the result was statistically significant.

- Two out the three attitudinal variables also have clear sign. Those who disagreed with *Ought to stay at home if below normal performance (yes=0)* and *Wrong to go to work if I infect colleagues, customers (yes=0)* also had higher presenteeism than those who agreed. The latter statement usually was significant.

- The final consistent result across the models is *sickness absence*. There is a positive and significant relationship to days with presenteeism.

In the two tables below the regression results are related to the 5 blocks of explanatory variables mentioned on p.21, and the 8 hypotheses, p. 21-22.

Block of variables	*Results of analysis (judged in terms of significant coefficients at 5% level)*
Health status	Very important across regression models
Socio-demographics	Age not important, age-squared only significant once. Gender is important, i.e. females have less presenteeism than men
Attitudinal variables:	No consistent picture for the three attitudinal variables. 'Scattered' significant coefficients. If 'fear of unemployment' is classified as attitudinal, a very clear picture emerges: Systematically significant.
Type of job and characteristics of the job	Managers consistently have lower presenteeism than non-managers. – Commitment to work is also important while satisfaction with place of work is not important. Work environment/type of task is rarely important. Exception: 'others handle my task if absent
Type of company and company policies:	Size in terms of no. of employees is not important. Presenteeism is higher in public than private companies – Type of remuneration is not important.

The 8 hypotheses are rejected in 3 out of 7 possible cases. When reading the table also read to comments to the hypotheses (p. 21-22). The terminology is ‘positivistic’, i.e. hypotheses can be rejected but not confirmed.

Hypothesis 1: When cost of absence rises, absenteeism decreases and presenteeism increases.	Rejected based on type of remuneration
Hypothesis 2: Absenteeism (resp. presenteeism) is lower (resp. higher) for employees with ‘team responsibility’.	Rejected
Hypothesis 3: As hierarchical level rises, absenteeism decreases and presenteeism increases.	Not rejected based on management responsibility
Hypothesis 4: Job insecurity associated with fixed-term contract deters absenteeism and reinforces presenteeism.	Not rejected based on ‘fear of unemployment’.
Hypothesis 5: Both absenteeism and presenteeism could be affected by a job mobility.	No data available
Hypothesis 6: Both absence and presenteeism are (slightly) correlated with children at home.	(based on separate regressions where number of children at home was explanatory variable): Rejects – apart from the logit part of hurdle model where coefficient was positive and significant.
Hypothesis 7: Both absenteeism and presenteeism are gender dependent.	Not rejected
Hypothesis 8: Both absence and presenteeism are age dependent	Rejected

.

Table 5: Negative binominal regression of determinants of presenteeism. Robust std. Stata routine: nbreg.

Dep. var: presenteeism days past three months, figure 2	**Coef.**	**Std. Err.**	**z**	**P>z**
Health Status				
No of chronic illnesses ((max of 14)	.0685801	.028316	2.42	**0.015**
Self rated health (excellent=0)				
* Good	.3231849	.1030916	3.13	**0.002**
* Acceptable	.5588752	.1167683	4.79	**0.000**
* Bad & really bad	1.098.222	.1980728	5.54	**0.000**
Socio-demographics				
Age	.0337158	.0214467	1.57	0.116
Age squared	-.0005385	.0002513	-2.14	**0.032**
Gender (male=0)	-.1651342	.0691652	-2.39	**0.017**
Occupation (self employed =0)				
* Skilled	-.2265603	.2556691	-0.89	0.376
* Unskilled	-.1683034	.2493525	-0.67	0.500
* 'white collar'	-.4937538	.2136942	-2.31	**0.021**
* Not classified	-.2716842	.2605059	-1.04	0.297
Attitude to absence/presenteeism				
Ought to stay at home if below normal performance (yes=0)				
* neither/nor - neutral	.2042354	.0839533	2.43	**0.015**
* No, disagrees	.1007794	.0802461	1.26	0.209
Wrong to go to work if I infect colleagues, customers (yes=0)				
* neither/nor - neutral	.1752118	.0980377	1.79	0.074
* No, disagrees	.2929173	.1280504	2.29	**0.022**
By calling in sick you recuperat faster (yes==)				
* neither/nor - neutral	-.10643	.0862217	-1.23	0.217
* No, disagrees	-.0119662	.1128425	-0.11	0.916
Fear of unemployment (to a very high degree =0)				
* some degree	-.3966877	.1808422	-2.19	**0.028**
* lesser degree	-.5118179	.1603595	-3.19	**0.001**
* not at all	-.4456435	.1521153	-2.93	**0.003**
The work place				
No mgt. Responsibility (0=mgt responsibility)	-.2390673	.0788686	-3.03	**0.002**
Commitment and involvement of work (very high=0)				
* high	.1287506	.1232706	1.04	0.296
* to some degree	.0325997	.1072949	0.30	0.761
* lesser degree	-.2505672	.1148531	-2.18	**0.029**

* not at all	-.6071554	.1416568	-4.29	**0.000**
Type of company (private=0)				
* Public	.1536317	.0750324	2.05	**0.041**
Seniority, no. years	-.0017659	.0045938	-0.38	0.701
No employees,(1-20=0)				
* 20-99 employees	.133448	.0941257	1.42	0.156
* 100-499 employees	.162712	.1010564	1.61	0.107
* > 499 employees	.1032522	.0996225	1.04	0.300
Sickness absence interview (yes=0)	.026421	.1024466	0.26	0.796
Satisfaction with place of work (very =0)				
* rather satisfied	-.021708	.0800382	-0.27	0.786
* satisfied	.1305811	.1010771	1.29	0.196
* dissatisfied	.0076841	.1324769	0.06	0.954
* rather and very dissatisfied	.1076158	.1677277	0.64	0.521
Type of work/task environment ('always or often'=0')				
* other handle my tasks if absent	.1642985	.0792444	2.07	**0.038**
* plan your own work	.066598	.0857167	0.78	0.437
***** physically exhausing work	-.0274104	.0929559	-0.29	0.768
Renumeration (fix monthly=0)				
* hourly	-.1117277	.1341038	-0.83	0.405
* base plus piece or commission	-.1249216	.1495781	-0.84	0.404
* mixture	.1441972	.1791456	0.80	0.421
Sickness absence, days past 12 months	.0255789	.004049	6.32	**0.000**
Constant	.5189579	.5135267	1.01	0.312
alpha	.8569428	.1123102		
N=3507				

nbreg in StatA fits two different parameterizations of the negative binomial model. The default, given by the dispersion(mean) option, has dispersion for the ith observation equal to 1 + alpha*exp(x_jb + offset_j); that is, the dispersion is a function of the expected mean of the counts for the jth observation. The alternative parameterization, given by the dispersion(constant) option, has dispersion equal to 1 + delta; that is, it is a constant for. Here (table 5) the default has been used. - Both parameterizations will yield similar results, and the parameterizations will usually not significantly differ from each other. Hence, the choice of parameterization is usually not important.

Table 6: Zero-inflated negative binomial regression, ZINB. Robust std er. Part 2. Stata routine:zinb

Dep. var: presenteeism days past three months, figure 2	**Coef.**	**Std. Err.**	**z**	**P>z**
Health Status				
No of chronic illnesses ((max of 14)	.0504672	.0314528	1.60	0.109
Self rated health (excellent=0)				
* Good	.1301397	.1064275	1.22	0.221
* Acceptable	.3127203	.1237578	2.53	**0.012**
* Bad & really bad	.7327003	.170997	4.28	**0.000**
Socio-demographics				
Age	.0351903	.0226206	1.56	0.120
Age squared	-.0004152	.0002628	-1.58	0.114
Gender (male=0)	-.180031	.0610862	-2.95	**0.003**
Occupation (self employed =0)				
* Skilled	-.2654493	.1823149	-1.46	0.145
* Unskilled	-.1916907	.1895514	-1.01	0.312
* 'white collar'	-.5180209	.1511007	-3.43	**0.001**
* Not classified	-.3628929	.1984816	-1.83	0.067
Attitude to absence/presenteeism				
Ought to stay at home if below normal performance (yes=0)				
* neither/nor - neutral	.1877846	.0796875	2.36	**0.018**
* No, disagrees	.0954269	.0732428	1.30	0.193
Wrong to go to work if I infect colleagues, customers (yes=0)				
* neither/nor - neutral	.1729375	.0854357	2.02	**0.043**
* No, disagrees	.2746212	.115368	2.38	**0.017**
By calling in sick you recuperat faster (yes==)				
* neither/nor - neutral	-.0948724	.0741697	-1.28	0.201
* No, disagrees	-.0114607	.096372	-0.12	0.905
Fear of unemployment (to a very high degree =0)				
* some degree	-.4147133	.1558685	-2.66	**0.008**
* lesser degree	-.4848414	.1310011	-3.70	**0.000**
* not at all	-.4312436	.1200831	-3.59	**0.000**
The work place				
No mgt. Responsibility (0=mgt responsibility)	-.2518902	.071587	-3.52	**0.000**
Commitment and involvement of work (very high=0)				
* high	.1211522	.1255906	0.96	0.335
* to some degree	.0660114	.1179766	0.56	0.576
* lesser degree	-.1965708	.121317	-1.62	0.105
* not at all	-.5279369	.1419145	-3.72	**0.000**
Type of company (private=0)				

* Public	.1613458	.0634363	2.54	**0.011**
Seniority, no. years	-.0006996	.0039889	-0.18	0.861
No employees,(1-20=0)				
* 20-99 employees	.1025441	.0833355	1.23	0.219
* 100-499 employees	.1426222	.0916715	1.56	0.120
* > 499 employees	.1017616	.0898844	1.13	0.258
Sickness absence interview (yes=0)	.0463726	.0907236	0.51	0.609
Satisfaction with place of work (very =0)				
* rather satisfied	-.0245576	.0763323	-0.32	0.748
* satisfied	.1327346	.084509	1.57	0.116
* dissatisfied	.0051447	.1350168	0.04	0.970
* rather and very dissatisfied	.1442304	.1644101	0.88	0.380
Type of work/task environment ('always or often'=0')				
* other handle my tasks if absent	-.0158155	.0789089	-0.20	0.841
* plan your own work	-.1054463	.1043858	-1.01	0.312
***** physically exhausing work	-.1424265	.1636381	-0.87	0.384
Renumeration (fix monthly=0)				
* hourly	-.1054463	.1043858	-1.01	0.312
* base plus piece or commission	-.1424265	.1636381	-0.87	0.384
* mixture	.1320915	.1501816	0.88	0.379
Sickness absence, days past 12 months	.0262955	.0031897	8.24	**0.000**
Constant	.5531603	.5398345	1.02	0.306
alpha	1.790.265	.1620741		
N=3507 of which 1762 are non-zero				

Table 7: Zero-inflated negative binomial regression, ZINB, Robust std err : Part 1 (inflated part, logit)

Health Status	Coef.	Std. Err.	z	P>z
No of chronic illnesses ((max of 14)	-.0906836	.1343031	-0.68	0.500
Self rated health (excellent=0)				
* Good	-1.168.592	.3132194	-3.73	**0.000**
* Acceptable	-1.632.274	.4332071	-3.77	**0.000**
* Bad & really bad	-2.779.803	1.350.766	-2.06	**0.040**
Socio-demographics				
Age	.3351164	.2781594	1.20	0.228
Age squared	-.0025787	.0026057	-0.99	0.322
Constant	-1.036.236	7.420.845	-1.40	0.163

In a zero-inflated model it is assumed that zero outcome is due to two different processes. For presenteeism the two processes are that a person has presenteeism or not. If not presenteeism only outcome possible is zero. If presenteeism it is then a count process. The two parts of the a zero-inflated model are a binary model, usually a logit model to model which of the two processes the zero outcome is associated with and a count model, in this case, a negative binomial model, to model the count process.

The Vuong test compares the zero-inflated model negative binomial with an ordinary negative binomial regression model. A significant z-test indicates that the zero-inflated model is preferred. This is clearly indicated in this analysis: Vuong test statistics: $z=3.59$, $Pr > z = 0.002$.

The likelihood ratio test that alpha = 0 is significantly different from zero, suggesting that the data are overdispersed and that a zero-inflated negative binomial model is more appropriate than a negative binomial model, hence supporting the Vuong test statistics.

Table 8: Two part model of presenteeism, part 1: Logit; part 2 negative binomial. Robust std err. Stata routine: tmp

Dep. var: presenteeism days past three months, figure 2	**Part 1 logit**			**Part 2: Binomial**		
Health Status	Coef.	Std. Err.	P>z	**Coef.**	Std. err.	P>z
No of chronic illnesses ((max of 14)	.097242	.0355831	**0.006**	.0232365	.0245172	0.343
Self rated health (excellent=0)						
* Good	.4204026	.103795	**0.000**	.095887	.0900186	0.287
* Acceptable	.6853385	.1268076	**0.000**	.233449	.1028228	**0.023**
* Bad & really bad	1.082.116	.2227825	**0.000**	.5442245	.1443333	**0.000**
Socio-demographics						
Age	.0150363	.0261761	0.566	.0211818	.0201048	0.292
Age squared	-.00038	.0003003	0.206	-.0002728	.000232	0.239
Gender (male=0)	-.182179	.0755787	**0.016**	-.0829803	.0577037	0.150
Occupation (self employed =0)						
* Skilled	.3974676	.223859	0.076	-.4356189	.1699166	**0.010**
* Unskilled	.2670741	.232862	0.251	-.2909505	.1792818	0.105
* 'white collar'	.2395573	.1823659	0.189	-.5821133	.1403916	**0.000**
* Not classified	.3303733	.2444145	0.176	-.4479109	.1877659	**0.017**
Attitude to absence/presenteeism						
Ought to stay at home if below normal performance (yes=0)						
* neither/nor - neutral	.1649466	.0961867	0.086	.1219724	.0753557	0.106
* No, disagrees	.1655521	.0902699	0.067	.0386406	.0693785	0.578
Wrong to go to work if I infect colleagues, customers (yes=0)						
* neither/nor - neutral	.179483	.1047257	0.087	.0723701	.0802927	0.367
* No, disagrees	.3002054	.1453538	**0.039**	.1394892	.1060869	0.189
By calling in sick you recuperat faster (yes==)						
* neither/nor - neutral	-.2177805	.0900571	**0.016**	.0195001	.0706416	0.783
* No, disagrees	-.2449652	.1166745	**0.036**	.0916307	.0907714	0.313
Fear of unemployment (to a very high degree =0)						
* some degree	-.3281119	.2090039	0.116	-.3243708	.1398423	**0.020**
* lesser degree	-.5411107	.1763826	**0.002**	-.3103441	.1174037	**0.008**
* not at all	-.4206036	.1648675	**0.011**	-.3038524	.1061709	**0.004**
The work place						
No mgt. Responsibility (0=mgt responsibility)	-.082532	.0879643	0.348	-.2044368	.0668013	0.002
Commitment and involvement of work (very high=0)						
* high	-.1030956	.1704456	0.545	.1252952	.1112277	0.260
* to some degree	-.3402686	.1578067	**0.031**	.1632488	.1049288	0.120
* lesser degree	-.6698065	.160358	**0.000**	.0536451	.1092642	0.623
* not at all	-1.166.083	.1810299	**0.000**	-.0104164	.1358523	0.939

Type of company (private=0)						
* Public	.0920969	.0797351	0.248	.1254007	.0600481	**0.037**
Seniority, no. years	-.0116817	.0046196	**0.011**	.0026866	.0037453	0.473
No employees,(1-20=0)						
* 20-99 employees	.06296	.1041375	0.545	.0811804	.0786693	0.302
* 100-499 employees	.1047219	.1125719	0.352	.0939585	.0861205	0.275
* > 499 employees	-.0645242	.1107011	0.560	.1431464	.0855798	0.094
Sickness absence interview (yes=0)	-.1367214	.1175183	0.245	.1053995	.083231	0.205
Satisfaction with place of work (very =0)						
* rather satisfied	.0711037	.0914936	0.437	-.0635835	.0727976	0.382
* satisfied	.0535092	.1027417	0.602	.0894867	.0803848	0.266
* dissatisfied	.1198842	.1700206	0.481	-.0513644	.1260605	0.684
* rather and very dissatisfied	-.0851073	.2097513	0.685	.1401277	.1551076	0.366
Type of work/task environment ('always or often'=0')						
* other handle my tasks if absent	.1444049	.0888007	0.104	.0700209	.0690254	0.310
* plan your own work	-.0430863	.0915542	0.638	.1059968	.0693497	0.126
* physically exhausing work	-.015129	.1014596	0.881	-.0052211	.0741166	0.944
Renumeration (fix monthly=0)						
* hourly	.0322698	.1323738	0.807	-.1107232	.1011994	0.274
* base plus piece or commission	.1231845	.2090456	0.556	-.1901834	.1541771	0.217
* mixture	.3345398	.1816419	0.066	.0168655	.1365679	0.902
Sickness absence, days past 12 months	.0060473	.0036803	0.100	.0219307	.002619	**0.000**
Constant	.334993	.6382552	0.600	1.246.881	.4908143	0.011
alpha						
N=3507						

Table 9: Latent class (finite mixture model) of determinants of presenteeism. Stata routine: fmm

	COMPONENT 1			COMPONENT 2		
Dep. var: presenteeism days past three months, figure 2	Coef.	Std. Err.	P>z	Coef.	Std. Err.	P>z
Dep. var: presenteeism days past three months, figure 2						
Health Status						
No of chronic illnesses ((max of 14)	.0984364	.0503707	**0.051**	.0359338	.0510342	**0.481**
Self rated health (excellent=0)						
* Good	.1634496	.1119261	0.144	.7767582	.3971258	**0.050**
* Acceptable	.2588988	.1415766	0.067	1.381.667	.4167906	**0.001**
* Bad & really bad	.5480939	.2036164	**0.007**	1.495.066	.456096	**0.001**
Socio-demographics						
Age	-.0182822	.030258	0.546	.0771638	.0696586	0.268
Age squared	.0001065	.0003596	0.767	-.000991	.0008112	0.222
Gender (male=0)	-.1397914	.0746186	0.061	-.0988328	.1728508	0.567
Occupation (self employed =0)						
* Skilled	.8632876	.2932281	**0.003**	-1.029.478	.4558381	**0.024**
* Unskilled	.611856	.3171462	**0.054**	-.2032157	.4232587	0.631
* 'white collar'	.7523024	.2791878	**0.007**	-.9504166	.3909066	**0.015**
* Not classified	.9044357	.3085256	**0.003**	-112.115	.7221375	0.121
Attitude to absence/presenteeism						
Ought to stay at home if below normal performance (yes=0)						
* neither/nor - neutral	.1521532	.0970873	0.117	.2608359	.2699501	0.334
* No, disagrees	.094888	.0911932	0.298	.1799611	.2557189	0.482
Wrong to go to work if I infect colleagues, customers (yes=0)						
* neither/nor - neutral	.0833161	.0964068	0.387	.2388103	.2149651	0.267
* No, disagrees	.2694368	.135548	**0.047**	.3024186	.2587314	0.242
By calling in sick you recuperat faster (yes==)						
* neither/nor - neutral	-.0918647	.0879759	0.296	-.2362441	.2201989	0.283
* No, disagrees	-.2694583	.1314771	**0.040**	.1070669	.2619518	0.683
Fear of unemployment (to a very high degree =0)						
* some degree	-.1341454	.1708829	0.432	-.530667	.322115	0.099
* lesser degree	-.0488302	.1433285	0.733	-.9546205	.3268899	**0.003**
* not at all	-.0368007	.1298234	0.777	-.8189549	.2644178	**0.002**
The work place						
No mgt. Responsibility (0=mgt responsibility)	-.0168055	.0884695	0.849	-.3314547	.2133388	0.120

Commitment and involvement of work (very high=0)						
* high	-.0028019	.129524	0.983	-.1952778	.2679757	0.466
* to some degree	-.161048	.1268796	0.204	-.2235052	.2555937	0.382
* lesser degree	-.3318715	.1463537	**0.023**	-.7340277	.3034344	**0.016**
* not at all	-.4480904	.2227148	**0.044**	-201.008	.5501176	**0.000**
Type of company (private=0)						
* Public	-.0564838	.0771292	0.464	.4842401	.1942302	0.013
Seniority, no. years	-.0109249	.0065733	0.097	-.0004949	.0129377	0.969
No employees,(1-20=0)						
* 20-99 employees	.1090214	.0912567	0.232	.0085194	.2605943	0.974
* 100-499 employees	-.1014668	.1125767	0.367	.5038788	.2856584	0.078
* > 499 employees	-.0728134	.1125457	0.518	.2976572	.3067371	0.332
Sickness absence interview (yes=0)	-.2634261	.099825	**0.008**	.5892538	.2674711	**0.028**
Satisfaction with place of work (very =0)						
* rather satisfied	.0707086	.0961162	0.462	-.1015359	.2458274	0.680
* satisfied	.0969947	.1008495	0.336	-.0820574	.2265155	0.717
* dissatisfied	.1000199	.1522801	0.511	-.1508637	.3706193	0.684
* rather and very dissatisfied	.2467567	.2075612	0.235	-.4354163	.5396806	0.420
Type of work/task environment ('always or often'=0')						
* other handle my tasks if absent	.2303304	.0913848	**0.012**	-.2459145	.1769252	0.165
* plan your own work	-.0732663	.1001992	0.465	.2510599	.2047284	0.220
* physically exhausing work	-.1142249	.0957493	0.233	.1478339	.2296355	0.520
Renumeration (fix monthly=0)						
* hourly	-.2109274	.1500227	0.160	.2706478	.2904237	0.351
* base plus piece or commission	-.037143	.1800347	0.837	.3131991	.3756239	0.404
* mixture	.1216635	.1532364	0.427	.1703373	.283537	0.548
Sickness absence, days past 12 months	-.0094581	.0031035	**0.002**	.0444735	.0070956	**0.000**
Constant	.8883607	.6932322	0.200	-.6132379	1.633.102	0.707

As noted above Deb and Trivedi hypothesized that the underlying unobserved heterogeneity which splits the population into latent classes is based on an individual's latent long-term health status for the demand for health care. This, however, may also be a possible interpretation for presenteeism. Proxy variables such as self-perceived health status and chronic health conditions may not fully capture population heterogeneity from this source. Consequently, in the case of two latent subpopulations, a distinction may be made between the "healthy" and the "ill" groups characterized by low mean and high mean, respectively. Hence, no sharp distinction is made 'no presenteeism' and 'positive presenteeism' as in the hurdle model. For a somewhat different interpretation see the discussion around figur5 below.

Table 10: Summary of coefficient signs and statistical significance (<0.05) across statistical models

					Two part model				**Latent class**			
	negativ bino		*ZINB*		*part 1 logit*		*Part 2 binom*		*Comp. 1*		*Comp2*	
	sign	signific	sign	signific	sign	signific	sign	signific	sign	signific	sign	signific
Dep. var: presenteeism days past three months, figure 2												
Health Status												
No of chronic illnesses ((max of 14)	+	**0.015**	+	0.109	+	**0.006**	+	0.343	+	**0.051**	+	**0.481**
Self rated health (excellent=0)												
* Good	+	**0.002**	+	0.221	+	**0.000**	+	0.287	+	0.144	+	**0.050**
* Acceptable	+	**0.000**	+	**0.012**	+	**0.000**	+	**0.023**	+	0.067	+	**0.001**
* Bad & really bad	+	**0.000**	+	**0.000**	+	**0.000**	+	**0.000**	+	**0.007**	+	**0.001**
Socio-demographics												
Age	+	0.116	+	0.120	+	0.566	+	0.292	-	0.546	+	0.268
Age squared	-	**0.032**	-	0.114	-	0.206	-	0.239	+	0.767	+	0.222
Gender (male=0)	-	**0.017**	-	**0.003**	-	**0.016**	-	0.150	-	0.061	+	0.567
Occupation (self employed =0)												
* Skilled	-	0.376	-	0.145	+	0.076	-	**0.010**	+	**0.003**	+	**0.024**
* Unskilled	-	0.500	-	0.312	+	0.251	-	0.105	+	**0.054**	+	0.631
* 'white collar'	-	**0.021**	-	**0.001**	+	0.189	-	**0.000**	+	**0.007**	+	**0.015**
* Not classified	-	0.297	-	0.067	+	0.176	-	**0.017**	+	**0.003**	+	0.121
Attitude to absence/presenteeism												
Ought to stay at home if below normal performance (yes=0)												
* neither/nor - neutral	+	**0.015**	+	**0.018**	+	0.086	+	0.106	+	0.117	+	0.334
* No, disagrees	+	0.209	+	0.193	+	0.067	+	0.578	+	0.298	+	0.482
Wrong to go to work if I infect colleagues, customers (yes=0)												
* neither/nor - neutral	+	0.074	+	**0.043**	+	0.087	+	0.367	+	0.387	+	0.267
* No, disagrees	+	**0.022**	+	**0.017**	+	**0.039**	+	0.189	+	**0.047**	+	0.242
By calling in sick you recuperat faster (yes==)												
* neither/nor - neutral	-	0.217	-	0.201	-	**0.016**	+	0.783	-	0.296	+	0.283
* No, disagrees	-	0.916	-	0.905	-	**0.036**	+	0.313	-	**0.040**	+	0.683
Fear of unemployment (to a very high degree =0)												
* some degree	-	**0.028**	-	**0.008**	-	0.116	-	**0.020**	-	0.432	+	0.099
* lesser degree	-	**0.001**	-	**0.000**	-	**0.002**	-	**0.008**	-	0.733	+	**0.003**
* not at all	-	**0.003**	-	**0.000**	-	**0.011**	-	**0.004**	-	0.777	+	**0.002**
The work place												
No mgt. Responsibility (0=mgt responsibility)	-	**0.002**	-	**0.000**	-	0.348	-	0.002	-	0.849	+	0.120
Commitment and involvement of work (very high=0)												

* high	+	0.296	+	0.335	-	0.545	+	0.260	-	0.983	-	0.466
* to some degree	+	0.761	+	0.576	-	**0.031**	+	0.120	-	0.204	-	0.382
* lesser degree	-	**0.029**	-	0.105	-	**0.000**	+	0.623	-	**0.023**	-	**0.016**
* not at all	-	**0.000**	-	**0.000**	-	**0.000**	-	0.939	-	**0.044**	-	**0.000**
Type of company (private=0)												
* Public	+	**0.041**	+	**0.011**	+	0.248	+	**0.037**	-	0.464	+	0.013
Seniority, no. years	-	0.701	-	0.861	-	**0.011**	+	0.473	-	0.097	-	0.969
No employees,(1-20=0)												
* 20-99 employees	+	0.156	+	0.219	+	0.545	+	0.302	+	0.232	+	0.974
* 100-499 employees	+	0.107	+	0.120	+	0.352	+	0.275	-	0.367	+	0.078
* > 499 - employees	+	0.300	+	0.258	-	0.560	+	0.094	-	0.518	+	0.332
Sickness absence interview (yes=0)	+	0.796	+	0.609	-	0.245	+	0.205	-	**0.008**	+	**0.028**
Satisfaction with place of work (very =0)												
* rather satisfied	-	0.786	-	0.748	+	0.437	-	0.382	+	0.462	-	0.680
* satisfied	+	0.196	+	0.116	+	0.602	+	0.266	+	0.336	-	0.717
* dissatisfied	+	0.954	+	0.970	+	0.481	-	0.684	+	0.511	-	0.684
* rather and very dissatisfied	+	0.521	+	0.380	-	0.685	+	0.366	+	0.235	-	0.420
Type of work/task environment ('always or often'=0')												
* other handle my tasks if absent	+	**0.038**	-	0.841	+	0.104	+	0.310	+	**0.012**	-	0.165
* plan your own work	+	0.437	-	0.312	-	0.638	+	0.126	-	0.465	+	0.220
* physically exhausing work	-	0.768	-	0.384	-	0.881	-	0.944	-	0.233	+	0.520
Renumeration (fix monthly=0)												
* hourly	-	0.405	-	0.312	+	0.807	-	0.274	-	0.160	+	0.351
* base plus piece or commission	-	0.404	-	0.384	+	0.556	-	0.217	-	0.837	+	0.404
* mixture	+	0.421	+	0.379	+	0.066	+	0.902	+	0.427	+	0.548
Sickness absence, days past 12 months	+	**0.000**	+	**0.000**	+	0.100	+	**0.000**	-	**0.002**	+	**0.000**
Constant		0.312		0.306		0.600		0.011		0.200		0.707

Model evaluation and testing

The following count models have been used: 1. Negative binominal regression, NB (table 5), 2. zero inflated negative binominal (ZINB), 3. the two part model, TMP, and 4. the latent variable (finite mixture models). The question is which of these models is the 'best'? It is a statistical question, but important because – as seen in the section on results – the count models vary in terms of which variables are significant. The approach followed here is similar to that of Deb & Trivedi and Cameron and Trivedi (chapter 6)[56, 64].

Nested – non-nested. Two models are non-nested models if neither model can be represented as a special case of the other. The likelihood ratio test, LR, can be used to discriminate between the models. The LR has the chi-square distribution, $\chi^2(p)$ where p is the difference in the number of parameters in the model.

The negative binomial models vs. Zinb has already been addressed above: The Vuong test compares the zero-inflated model negative binomial with an ordinary negative binomial regression model. A significant z-test indicates that the zero-inflated model is preferred. This is clearly indicated in this analysis: Vuong test statistics: z=3.59, Pr > z = 0.002.

The likelihood ratio test that alpha = 0 is significantly different from zero, suggesting that the data are overdispersed and that a zero-inflated negative binomial model is more appropriate than a negative binomial model

It is standard to use information criteria to compare non-nested models. The Akaike information criterion, AIC, and the Bayesian information criterion, BIC, are the two standard possibilities

The AIC and the BIC are two popular measures for comparing maximum likelihood models. AIC and BIC are defined as AIC = -2*ln(likelihood) + 2*k, BIC = -2*ln(likelihood) + ln(N)*k where k = number of parameters estimated and N = number of observations

Suggested strategy for model selection:

1. Use AIC and BIC to compare finite mixture models with varying latent classes. Here, however, only a two-class model is used.
2. Use Likelihood ratio tests to compare negative binomial model with two-part model
3. Use AIC and BIC to compare finite mixture models with two part model.

AIC is an estimate of a constant plus the relative distance between the unknown true likelihood function of the data and the fitted likelihood function of the model, so that a lower AIC means a model is considered to be closer to the truth. BIC is an estimate of a function of the posterior probability of a model being true, under a certain Bayesian setup, so that a lower BIC means that a model is considered to be more likely to be the true model.

Table 11: Test statistics for model selection

	k	lnL	AIC	BIC
Vuong test of zinb vs. standard negative binomial: z = 3.59 Pr>z = 0.0002				
ZINB (two part)	48	-6346	12794	**13108**
Hurdle/two part model (logit and negative binomial 2 (means dispersion)	121	-6297	12836	13582
Finite Mixture, 2 components - negative binomial 2 (means dispersion)	89	-6260	**12698**	**13246**

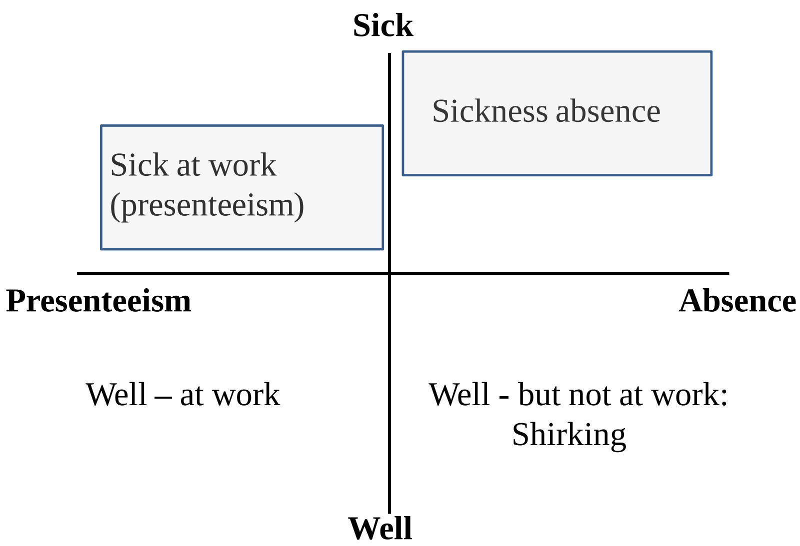
The comparison of the two-part model ('hurdle') and finite mixture 2-component points towards the finite mixture model as the preferred one both for AIC and BIC while the ZINB is the preferred if judged alone on BIC, and ZINB is preferred to the standard negative binomial model (Vuong test).

Overall it appears that the finite mixture model ought to be the preferred model based on statistical criteria along with a relevant interpretation.

Hence, the regression results in table 9 ought to attract attention. As to the interpretation of the two latent classes (see text below table 9) Deb and Trivedi hypothesized for the case of demand for health services that the underlying unobserved heterogeneity which splits the population into latent classes is based on an individual's latent long-term health status which also happens to be relevant for presenteeism. Proxy variables such as self-perceived health status and chronic health conditions may not fully capture population heterogeneity from this source. Consequently, in the case of two latent subpopulations, a distinction may be made between the "healthy" and the "ill" groups, whose demands for medical care (or sickness absence/presenteeism) are characterized by low mean and high mean, respectively.

However, there may be an alternative interpretation based on figure 5, namely a subjective threshold-thinking as a source of heterogeneity generating two (latent) classes. The latent threshold then concerns the tendency to report sick/go sick to work where one can easily imagine considerable heterogeneity generating (at least) two distinct classes with different threshold values.

Figure 5: Classification



Conclusions and perspectives

From an economic viewpoint the modeling of presenteeism is still in its infancy both for the determinants of productivity (inclusive of valuation of productivity losses) and the determinants of the extent (volume) of presenteeism measured in terms of days with presenteeism.

As regards the latter only two theoretical models were identified and found lacking in part because the essentials of presenteeism have not been captured. However, the rather general approach of discrete choice (random utility) seems to be an appropriate approach that combines utility maximizing behavior with the econometrics of count models. This apparently is the first time this approach has been used for the understanding and analysis of presenteeism data.

The other economic dimension, namely the question of possibly decreased productivity in connection with presenteeism is well researched, still with relevant issues surrounding the 'right' monetary valuation of the possible productivity loss. A far more intriguing question concerns the inclusion of indirect costs in economic evaluations. However, this is left to experts in this particular field of health economics to discuss.

In terms of econometrics the use of count models and two part models for the analysis of presenteeism data is still in its infancy. For instance, the use of the two part model and latent variable model has to our knowledge only been attempted so far in the present paper. In view of the grey area around sickness absence and presenteeism and the underlying health situation the ideas behind the latent variable approach seem relevant to explore further where the latency issue is a possible threshold where sickness absence sets in instead of presenteeism.

The empirical results for productivity loss of presenteeism do not point to considerable losses. For instance, very few reported postponing tasks even though that the work pace was slowed down. The present data of course do not provide evidence on possible negative medium and long term effects on health. In this connection one should remember that much presenteeism is related to persons with a chronic condition which may not necessarily worsen as a consequence of presenteeism. Rather, this group may have learned to cope with the situation. Presenteeism also can have root in a brief acute episode, e.g. a cold or a mild flu. It is unlikely that will result in serious permanent health impact, but the possible contagious effects vis-a-vis other employees or customers may actually be the worst effect.

As to the determinants of presenteeism the analyses across regression models point clearly towards the importance of:

- health status
- age and gender
- fear of unemployment
- attitude towards sickness absence/presenteeism

- the positive relationship to sickness absence
- having management responsibility

At present the use of no less than four count-related models seems to point towards a two-class latent finite mixture model as the preferred.

As noted the discrete choice/random utility is a very promising approach. However, in the model it is assumed that all choices are based on utility maximizing behavior. It raises the following question as posed by McFadden: "The problem of revealed stochastic preference asks the question: Are the distributions of choices observed for a population of individuals in a variety of choice situations consistent with rational choice theory, which postulates that individuals maximize preferences?"[65]. In other words, when can observed choice probabilities be rationalized as consistent with random utility maximization (RUM)? - Here we are not going into the rather advanced mathematical details of this question. However, it is important for a consistent theoretical base for the RUM-models – and all too rarely is this question posed. In addition, it is equally important to be able to test for the assumed utility maximizing behavior. This area seems to be in its infancy[66].

Reference List

1. Johns G. Presenteeism in the workplace: A review and research agenda. Journal of Organizatonal Behavior 2010;31(4):519-542.

2. Skåtun D, Skåtun JD. The Impact of Turnout on Turning Up: The Complementarity of Attendance among Co-Workers. Economic Journal 2004;83(3):225-242.

3. Brouwer WB, Koopmanschap MA, Rutten FF. Productivity losses without absence: measurement validation and empirical evidence. Health Policy 1999;48(1):13-27.

4. Brouwer WB, van Exel NJ, Koopmanschap MA, Rutten FF. Productivity costs before and after absence from work: as important as common? Health Policy 2002;61(2):173-187.

5. Koopmanschap M, Burdorf A, Jacob K, Meerding WJ, Brouwer W, Severens H. Measuring productivity changes in economic evaluation: setting the research agenda. Pharmacoeconomics 2005;23(1):47-54.

6. Brown S, Sessions JG. Absenteeism, 'Presenteeism', and Shirking. Economic Issues 2004;9(1):15-21.

7. Chatterji M, Tilley CJ. Sickness, Absenteeism, Presenteeism, and Sick Pay. Oxford Economic Papers 2002;54(4):669-687.

8. Pauly MV, Nicholson S, Xu J et al. A general model of the impact of absenteeism on employers and employees. Health Econ 2002;11(3):221-231.

9. Pauly MV, Nicholson S, Polsky D, Berger ML, Sharda C. Valuing reductions in on-the-job illness: 'presenteeism' from managerial and economic perspectives. Health Econ 2008;17(4):469-485.

10. Loeppke R, Taitel M, Haufle V, Parry T, Kessler RC, Jinnett K. Health and productivity as a business strategy: a multiemployer study. J Occup Environ Med 2009;51(4):411-428.

11. Mattke S, Balakrishnan A, Bergamo G, Newberry SJ. A review of methods to measure health-related productivity loss. Am J Manag Care 2007;13(4):211-217.

12. Prasad M, Wahlqvist P, Shikiar R, Shih YC. A review of self-report instruments measuring health-related work productivity: a patient-reported outcomes perspective. Pharmacoeconomics 2004;22(4):225-244.

13. Schultz AB, Chen CY, Edington DW. The Cost and Impact of Health Conditions on Presenteeism to Employers: A Review of the Literature. Pharmacoeconomics 2009;27(5):365.

14. Burton WN, Conti DJ, Chen CY, Schultz AB, Edington DW. The role of health risk factors and disease on worker productivity. J Occup Environ Med 1999;41(10):863-877.

15. Koopman C, Pelletier KR, Murray JF et al. Stanford presenteeism scale: health status and employee productivity. J Occup Environ Med 2002;44(1):14-20.

16. Lofland JH, Pizzi L, Frick KD. A review of health-related workplace productivity loss instruments. Pharmacoeconomics 2004;22(3):165-184.

17. Goetzel RZ, Long SR, Ozminkowski RJ, Hawkins K, Wang S, Lynch W. Health, absence, disability, and presenteeism cost estimates of certain physical and mental health conditions affecting U.S. employers. J Occup Environ Med 2004;46(4):398-412.

18. Koopmanschap MA, van Inevel BM. Towards a new approach for estimating indirect costs of disease. Soc Sci Med 34[9], 1005-1010. 1992.

19. Koopmanschap MA, Rutten FF. A practical guide for calculating indirect costs of disease. Pharmacoeconomics 1996;10:460-466.

20. Akobundu E, Ju J, Blatt L, Mullins CD. Cost-of-Illness Studies: A Review of Current Methods. Pharmacoeconomics 24[9], 869. 2006. New Zealand, Adis International.

21. Drummond MF, Stoddart GL, Torrance GW, O'Brien. Methods for the economic evaluation of health care programmes, second edition. Oxford: Oxford University Press; 1997.

22. Aronsson G, Gustafsson K, Dallner M. Sick but yet at work. An empirical study of sickness presenteeism. J Epidemiol Community Health 2000;54(7):502-509.

23. Aronsson G, Gustafsson K. Sickness presenteeism: prevalence, attendance-pressure factors, and an outline of a model for research. J Occup Environ Med 2005;47(9):958-966.

24. Adelmann J, Sharfman M, Johnson R, et al. Impact of Oral Sumatriptan on Workplace Productivity, Health-Related Quality of Life, Healthcare Use, and Patient Satisfaction with Medication in Nurses with Migraine . American Jouranl of Managed care 1996;2(11):1407-1416.

25. Osterhaus JT, Gutterman DL, Plachetka JR. Healthcare resource and lost labour costs of migraine headache in the US. Pharmacoeconomics 1992;2(1):67-76.

26. Gold MR, Siegel JE, Russell LB, Weinstein MC, (eds). Cost-effectiveness in health and medicine. Oxford & New York: Oxford University Press; 1996.

27. Koopmanschap MA, Rutten FFH, van Ineveld BM, van Roijen L. The friction cost method for measuring indirect costs of disease. Journal of Health Economics 1995;14:171-189.

28. Brown S, Sessions JG. The Economics of Absence: Theory and Evidence. Journal of Economic Surveys 1996;10:23-53.

29. Allen SG. An empirical model of worker attendance. Review of Economics and Statistics 1981;71(1):1-17.

30. Barmby T, Sessions JG, Treble JG. Absenteeism, Efficiency Wages and Shirking. Scandinavian Journal of Economics 1994;96(4):561-566.

31. Shapiro C, Stiglitz JE. Equilibrium Unemployment as a Worker Discipline Device: Reply. American Economic Review 1985;75(4):892-893.

32. Barmby T, Ercolani M, Treble J. Sickness Absence: An International Comparison. Economic Journal 2002;112:F315-F331.

33. Chatterji M, Tilley CJ. Sickness, Absenteeism, Presenteeism, and Sick Pay. Oxford Economic Papers 2002;54(4):669-687.

34. Granlund D. "Waiting for the other shoe to drop": Waiting for healtyh care and duration of sick leave. Working Paper 8/2010 Örebro University 2010.

35. Becker G. A theory of the allocation of time. The Economic Journal 1965;75(299):493-517.

36. Grossman M. The demand for health: a theoretical and empirical investigation. 1972.

37. Pedersen K. Sickness absence and voluntary employer paid health insurance. Odense: COHERE, 2011Health Economics Papers 2011: 1.)

38. Viscusi WK, Evans WN. Utility Functions That Depend on Health Status: Estimates and Economic Implications. American Economic Review80(3):353-374.

39. Finkelstein A, Luttmer E, Notowidigdo M. Approaches to Estimating the Health State Dependence of the Utility Function. American Economic Review: Papers and Proceeding 2009;99(2):116-121.

40. Ose SO. Working Conditions, Compensation and Absenteeism. Journal of Health Economics 2005;24(1):161-188.

41. Coles MG, Treble JG. The price of work reliablity. Economics Letters 1993;44:149-155.

42. Culyer A. The dictionary of health economics. Cheltenham: Edward Elgar; 2005.

43. Greene W. Econometric Analysis (7th edition). Upper Saddle River, New Jersey : Prentice Hall; 2011.

44. McFadden D. Economic Choices. The American Economic Review 2001;91(3):351-378.

45. Creedy J, Kalb G. Discrete Hours Labour Supply Modelling: Specification, Estimation and Simulation. Journal of Economic Surveys 2005;19(5):697-734.

46. Dagsvik JK, Jia Z, Kornstad T, Thoresen TO. Theoretical and Practical Arguments for Modeling Labor Supply as a Choice among Latent Jobs. CESifo Group Munich, CESifo Working Paper Series: 3708, 2012 . 2012. CESifo Group Munich, CESifo Working Paper Series: 3708.

47. Harding M, Burda M. A poisson mixture model of discrete choice. Journal of econometrics 2012;166(2):184-203.

48. Bockerman P, Laukkanen E. Predictors of sickness absence and presenteeism: does the pattern differ by a respondent's health? J Occup Environ Med 2010;52(3):332-335.

49. d'Uva TB. Latent class models for utilisation of health care. Health Econ 2006;15(4):329-343.

50. Deb P, Trivedi P. Empirical models of health care use. In: Jones A, editor. The Elgar Companion to Health Economics. Cheltenham: Edward Elgar; 2006:147-155.

51. Deb P. Demand for Medical Care by the Elderly: A Finite Mixture Approach. Journal of Applied Econometrics 1998;12:313-336.

52. Deb P, Trivedi PK. The structure of demand for health care: latent class versus two-part models. J Health Econ 2002;21(4):601-625.

53. Jimenez-Martin S, Labeaga JM, Martinez-Granado M. Latent class versus two-part models in the demand for physician services across the European Union. Health Econ 2002;11(4):301-321.

54. Bierla I, Huver B, Richard S. New evidence on absenteeism and presenteeism. The International Journal of HumanResource Management 2013;24(7):1536-1550.

55. Huver B, Richard S, Bierla I, Vaneecloo N, Delclite T. Sick but at Work.An Econometric Approach to Presenteeism. 2012.

Ref Type: Unpublished Work

56. Cameron A, Trivedi P. Regression Analysis of Count Data, second edition. New York: Cambridge University Press; 2013.

57. Rose CE, Martin SW, Wannemuehler KA, Plikaytis BD. On the use of zero-inflated and hurdle models for modeling vaccine adverse event count data. J Biopharm Stat 2006;16(4):463-481.

58. d'Uva TB. Latent class models for use of primary care: evidence from a British panel. Health Econ 2005;14(9):873-892.

59. d'Uva TB., Jones AM. Health care utilisation in Europe: new evidence from the ECHP. J Health Econ 2009;28(2):265-279.

60. Kessler RC, Barber C, Beck A et al. The World Health Organization Health and Work Performance Questionnaire (HPQ). J Occup Environ Med 2003;45(2):156-174.

61. Kessler RC, Ames M, Hymel PA et al. Using the World Health Organization Health and Work Performance Questionnaire (HPQ) to evaluate the indirect workplace costs of illness. J Occup Environ Med 2004;46(6 Suppl):S23-S37.

62. Turpin RS, Ozminkowski RJ, Sharda CE et al. Reliability and validity of the Stanford Presenteeism Scale. J Occup Environ Med 2004;46(11):1123-1133.

63. Meerding W, et al. Health problems lead to considerable productivity loss at work among workers with high physical load jobs. Journal of Clinical Epidemiology 2005;58(5):517-523.

64. Deb P, Trivedi PK. Demand for Medical Care by the Elderly: A Finite Mixture Approach. Journal of Applied Econometrics 1997;12(3):313-336.

65. McFadden D. Stochastic preference: a synthesis. Economic Theory 2005;26:245-264.

66. Kitamuary Y, toye JO. Nonparametric analysis of random utility models: testing . Working Paper Princeton 2012.