Overweight and Obesity and the Demand for Primary Physician Care

Nabanita Datta Gupta
Jane Greve

DISCUSSION PAPER SERIES

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Overweight and Obesity and the Demand for Primary Physician Care

Nabanita Datta Gupta
Aarhus School of Business
and IZA

Jane Greve
Rockwool Foundation Research Unit

ABSTRACT

Overweight and Obesity and the Demand for Primary Physician Care[*]

The standard economic model for the demand for health care predicts that unhealthy behaviour such as being overweight or obese should increase the demand for medical care, particularly as clinical studies link obesity to a number of serious diseases. In this paper, we investigate whether overweight or obese individuals demand more medical care than normal weight individuals by estimating a finite mixture model which splits the population into frequent and non-frequent users of primary physician (GP) services according to the individual's latent health status. Based on a sample of wage-earners aged 25-60 years drawn from the National Health Interview (NHI) survey 2000 and merged to Danish register data, we compare differences in the impact of being overweight and obese relative to being normal weight on the demand for primary physician care. Estimated bodyweight effects vary across latent classes and show that being obese or overweight does not increase the demand for primary physician care among infrequent users but does so among frequent users.

JEL Classification: I12, I18

Keywords: overweight, obesity, demand for primary physician care

Corresponding author:

Nabanita Datta Gupta
Department of Economics
Frichshuset Hermodsvej 22
8230 Aabyhøj
Denmark
E-mail: ndg@asb.dk

1. Introduction

The purpose of this paper is to empirically characterize and explain observed patterns of medical care utilization by overweight and obese people. Obesity is a precursor of serious diseases. The clinical evidence links obesity to a number of major illnesses including hypertension, stroke, type 2 diabetes, coronary artery disease (CAD), arthritis, gallbladder and liver disease, respiratory disorders, such as sleep apnea and asthma, and certain types of cancers (NIH, 1998). Thus, a clear testable prediction coming out of the standard economic model for the demand for health care would be that being overweight or obese should increase the demand for medical care compared to being normal weight. The aim of this paper is to test this prediction.

The motivation for examining the demand for medical care among the obese and overweight is the rapid increase in the costs of obesity related diseases world-wide. These costs include both the direct medical costs of diagnosis and treatment of diseases associated with obesity, and the indirect costs due to obesity related morbidity and mortality (Wolf and Colditz, 1998). Increased medical costs of obesity imply higher average taxes for all individuals, both obese and non-obese, in a health care system with public insurance provision. Even in a privately insured system, as long as premiums are not tied to weight, overweight individuals will pass on the costs of their treatment to the other insured in their insurance pool (Bhattacharya and Sood, 2005). In this way, obesity imposes an external cost on other individuals, and may constitute a rationale for government intervention (Grossman and Rashad, 2004; Lakdawalla and Philipson, 2006)[1].

In 2004, the US Health and Human Services declared obesity an illness which could be covered by Medicare (the US federal health insurance program for the elderly and disabled). This change meant that Medicare recipients could now submit their expenditures on approved treatments for reducing excessive weight for reimbursement. Such treatments range from enrollment in a weight-loss clinic or a fitness center to gastric bypass surgery[2]. 'Exercise on prescription' and other weight reducing

---

[1]However, a study which simulates lifetime medical expenses for a cohort of obese individuals aged 20 at baseline and compares it to similar cohorts of smokers and healthy persons finds that while annual health expenditures are highest for the group of obese up to the age of 56, beyond that age, smokers incur higher annual expenses. Furthermore, differences in life expectancies across these groups imply that normal weight people have the highest lifetime medical expenses, and smokers the lowest, with obese individuals falling in the middle (van Baal et al., 2008).
[2] Weight reducing drugs such as Orlistat and Sibutramine are not covered.

treatments are offered to the obese or overweight in many other countries as well (e.g., UK and Denmark). An evaluation of specific obesity reducing treatments is difficult to make because those obese individuals who initiate contacts to medical practitioners to take advantage of weight-reducing programs are likely to be non-randomly selected, i.e., they probably have more health knowledge and are the ones motivated to lose weight. Still, a number of studies show that obesity reducing treatments are associated with weight loss although there is no evidence that the effect is persistent (Svendsen et al., 2001).

Although it is difficult to treat obesity, treatment of obesity is always recommended when there are weight related complications. Medical researchers tend to agree that in a long run perspective preventing obesity is less costly than treating obesity – although this is still not known with certainty – and therefore they tend to recommend preventive treatments with focus on selected vulnerable groups as an instrument in a national plan for reducing rising obesity rates (Svendsen et al. 2001). A recent study suggests that as primary physicians (GP's) are easily accessible to the population, they could play a key role in the prevention of overweight and obesity by giving basic information on nutrition and lifestyle and by monitoring weight gain (Sculati et al., 2006).

In fact, so far, little is known about the intensity of use of medical care by overweight or obese individuals despite the fact that many countries are now prepared to allocate considerable resources to treating obesity as a public health problem which can best be taken care of by medical means. In this paper, we investigate whether overweight or obese individuals demand more primary physician care than normal weight individuals. However, instead of a single estimate of the demand for medical care we propose to extract information from the natural heterogeneity which exists in a population consisting of both healthy and unhealthy individuals distributed across the different weight classes. To do this we estimate a finite mixture variant of the latent class model which splits the population into frequent and non-frequent users of GP care according to the individual's latent health status. The findings from our study will inform whether overweight and obese individuals engage in preventive health care–e.g., by demanding medical care when they are healthy–or whether they tend to demand medical care only when unhealthy, i.e., when illnesses are diagnosed.

The rest of the paper is organized as follows: Section 2 gives some background on the literature on the demand for health care; Section 3 describes the data; Section 4 sets out the finite mixture model

and compares it to other empirical models proposed in the literature; Section 5 presents the results of the estimation; Section 6 discusses the policy implications of our findings and Section 7 concludes.

2. Background

The standard model for the demand for health and the demand for medical care is Grossman's (1972a and 1972b) model. In Grossman's model, wage-earning individuals maximize an inter-temporal utility function of commodities and health. Health is valued for itself, but also because good health affords individuals time which can be allotted to market and non-market activities. Individuals are endowed with an initial health stock which depreciates each period due to not only the ageing process but also to health related behaviors such as smoking, drinking and poor eating habits as well as influences of environmental conditions (Cropper, 1981). Individuals can, however, make replenishing investments each period through combining inputs such as medical care, diet, exercise etc., which allow them to choose their optimal health stock and ultimately their length of life. Demand for medical care is derived from the demand for health. Besides health status, demand for medical care depends on wage, age, education, prices for medical goods, and unhealthy behaviors such as smoking and obesity (Nocera and Zweifel, 1998; Wagstaff, 1986). A clear implication arising from Grossman's model is that unhealthy behavior such as being overweight or obese increases the demand for medical care.

A long list of studies finds that obesity is related to diseases such as hypertension, stroke, type 2 diabetes, coronary artery disease (CAD), arthritis, gallbladder and liver disease, respiratory disorders, such as sleep apnea and asthma, and certain types of cancers (NIH, 1998). These diseases require both frequent hospitalisation and frequent medical care. However, not all obese individuals have health problems. The medical literature distinguishes between people with and without the "metabolic syndrome", which is a clustering of several metabolic and cardiovascular disease risk factors (Isomaa, 2003; Madbad and Astrup, 2004). The metabolic syndrome includes lifestyle factors such as obesity, but not all obese individuals necessarily have it, and, some normal weight people could have it as well. Thus, unless diagnosed, some individuals (both normal and overweight) may be 'silent carriers' of obesity related diseases. Particularly, this may be the case

for young people who rarely get health care checkups done, but who make up the group which has experienced the fastest rise in obesity in recent times[3].

In this paper we look at a selected group (i.e. wage earners) of relatively healthy people in different weight categories to see whether people in higher weight categories who have a higher risk of lifestyle diseases use more primary physician care. A large epidemiological literature has established a clear positive association between BMI (or some other measure of overweight/obesity) and medical consumption. Most of these studies derive estimates of health care costs of obesity using 'population attributable risk' methods which relate the proportion of cases of a particular disease (diabetes, for example) to obesity (see, e.g., Ministry of Health and Prevention, 2007). A few studies are based on prospective cohort analysis which tracks the subsequent health care usage of a single cohort which is selected on the basis of exposure rather than disease. However, both types of analyses can at best uncover correlations and furthermore, are sensitive to selection or attrition problems. We are not aware of research which has directly examined the impact of being overweight and obese on the demand for medical care by employing a very general method which allows for varying intensities of medical care use on the basis of the frailty in the data.

Past work on behaviours and bodyweight may help to inform on the relationship between observed health status on the one hand, and health outcomes and behaviour on the other hand, as described below.

First, in the behavioural economics literature the phenomenon of time inconsistent preferences can be related to overweight and obesity if these states are the result of addictive behaviour. Time inconsistent preferences arise when people discount tomorrow relative to today more than they discount the day after tomorrow relative to tomorrow. Cawley (2006) describes how additive behaviour, such as eating and physical activity patterns, that lead to obesity can be a result of inconsistent time preferences. If increased bodyweight is related to inconsistent time preferences, then we will expect people in higher weight categories to postpone doctor visits unless they are forced to go to the doctor with, e.g., an acute disease. People with inconsistent time preferences know that going to the doctor would be best in the long run but keep putting such visits off until the

---

[3] According to Sanz-de-Galdeano (2005) Denmark ranks among the two least obese countries in Europe for men and women aged 45 or older. However, the same study also finds that among younger people (i.e., 15-to-29-year-olds), Denmark has one of the largest fractions of obese persons in the EU countries.

next day. Several survey studies have documented an explicit negative attitude among doctors and other health personnel to obese people (Hoppe and Odgen, 1997; Price et al., 1987; Puhl and Brownell, 2001). If overweight people experience or expect discrimination at the GP's office, they might begin to postpone a check up at the doctor. Another reason for not visiting the doctor more frequently may be the physical exertion that is required to make such a trip. Obese people may try to avoid the effort involved such as moving around and climbing stairs unless a specific reason— such as an acute disease—turns up. Thus, an activity which may be unpleasant for many reasons is postponed by individuals with time-inconsistent preferences.

Second, if people are less willing to trade current utility for potential future health benefits, then we expect them to spend less time on medical care. A number of studies have analyzed the relationship between bodyweight and preferences for time (Borghans and Golsteyn, 2005; Cutler and Glaeser, 2005; Komlos et al., 2004; Smith et al., 2005) and find some evidence of the relationship between increasing bodyweight and higher discounting of the future. But this relationship varies between specific sub-groups within the population and seems to depend a great deal on the proxy used for the discount rate.

Third, as mentioned previously, although obesity is associated with poor health, not all obese people are unhealthy (e.g., have the metabolic syndrome). People within the higher weight categories with low use of primary physician care might simply not have a greater demand for medical care than normal weight people because they are basically healthy and practice healthy behaviours.

In sum, although Grossman's model from 1972 predicts a positive relation between harmful health behaviour and demand for medical care, this might not be the case when we subdivide the sample into groups depending on health status and health behaviours. Both economic theory within behavioural economics and non-economic theory provide explanations as to why people with higher bodyweight would demand less physician care when they are relatively healthy. For that reason it is not clear how increased bodyweight is related to physician care.

3. Data description

3.1 Data and variable definitions

The data used in this study is combined survey (The National Health Interview survey, NHI) and register data. The NHI is a survey of individuals aged 16 years or more in 2000. The data collection was done via personal interviews which took place in the homes of the respondents. The NHI is a very rich survey with detailed information on both social and demographic variables, work environment, different types of physician care usage, self reported health, presence of chronic diseases and height and weight. Individuals were randomly sampled on a long list of variables such as sex, age, ethnic origin etc. However, there was over-sampling of individuals in small counties to ensure that a minimum of 1,000 individuals were interviewed in each county.[4] We control for county in all estimations to take into account this non-random selection.

We set up the empirical analogue to Grossman's model in which individuals face both time and money costs of investing in their health by limiting the analysis to wage-earners only (10,153 observations). Simple descriptive statistics shows that the unemployed and people out of the labor force use primary physician care much more than other socioeconomics groups (Statistics Denmark, 2000). By restricting the sample to wage earners only, we thereby minimize heterogeneity in the data with respect to doctor visits. Wage-earners between 25 and 60 years of age are selected. People below the age of 25 are excluded from the analysis to restrict the sample to full time (non-student) workers. As people start to withdraw from the labor force at age 60 in Denmark, people above this age are excluded. Omitting persons older than 60 and the non-employed means that, our sample is expected to be somewhat healthier and less obese than the overall population. Nonetheless, it is representative of the majority of prime-aged adults in Denmark.[5]

In this paper we look at differences in the demand for health care across different bodyweight categories for men and women pooled together.[6] Women require more health-care than men, particularly with respect to their reproductive lives. Furthermore, pregnancies are most often related to a significant weight gain. Therefore, women observed either when pregnant, or shortly after any

---

[4] Detailed information on data collection can be found in Ekholm et al. (2006).

[5] Although women who are full-time family carers also are included in the group of prime-aged adults in Denmark this group is left out of the analysis. However, very few women in Denmark are full-time family carers. In Denmark more than 80 percent of women in families with children under the age of 6 are employed (OECD, 2001).

[6] All models estimated in this paper have also been estimated separately by gender. However, when we estimate the FMM, too few observations are available in the gender specific estimations to give valid results.

child-birth event, are omitted (297 women are excluded because of that). Furthermore it is possible to identify women using contraception, which requires a yearly gynaecological examination, and define a control for this group in the analyses.

Observations are also excluded from the data because of missing information on GP visits, wage, or on other variables in the list of included survey variables. In total 117 observations are excluded because of missing observations. Out of these 117 observations, 81 have missing information on height and weight. To test whether the results are robust to the selection arising when people with missing height and weight are excluded from the analysis, all models are estimated with and without these 81 observations. When a dummy is included for people with missing information on bodyweight, the overall results do not change. Finally, 2 observations had an unusually high number of GP visits and were therefore excluded from the estimations. Consequently, the final number of observations used in this version of the paper is 8,036 wage earners.

Due to a unique identifier, the variables from NHI are merged with information from administrative registers in 2000 and 2005 which–besides containing thorough information on demographic measures, socioeconomic status, employment, unemployment and income–include information on usage of GP, hospitalizations, and bed days. As key explanatory variables such as self-reported height, weight, health, and smoking behavior only exist in 2000 for the full sample, register data from 2005 will only be used for out-of-sample validation.

The key outcome variable in this study is primary physician visits in 2000. Information on GP visits comes from the Danish Public Health Insurance register (*Sygesikringsregisteret*). It is through the public health insurance system that Danish citizens get free medical service. The main purpose of this register is to maintain a database for the purpose of reimbursing doctors on a fee-for-service basis. Therefore doctors in Denmark have incentives to report whether a treatment was given, the patient's social security number (CPR number), and the type of treatment correctly. Thus, this variable quite accurately records the use of medical care in the Danish population.

The second key component in the analysis is a measure for bodyweight. In this paper we use body mass index (BMI, measured as weight in kilograms over height in meters squared, kg/m^2), and different classifications of it. In particular, the focus is on people categorized as being overweight,

which means BMI between 25 and 30, and obese, which means BMI above 30. Obesity can be identified by other anthropometric measures such as skin-fold and waist circumference, but these measures are not available in the data. The category "normal weight" includes people with BMI below 25. This means that this category includes people with BMI below 18.5 as well, who are categorised as being underweight. In the sample 140 people (19 men and 121 women) are categorised as underweight. Leaving out these 140 observations on underweight individuals from the analysis does not change the estimation results.

3.2 Descriptive statistics

Table 1 presents a detailed variable description and Table 2 shows sample means and sample standard errors for all of the variables used in the analysis. The descriptive statistics show that on average wage-earners make 4.6 GP visits a year. The mean age in the sample is 43.

Table 2 shows that among wage-earners, 43 percent have a BMI above 25, which means that they can be categorized as being overweight or obese, 34 percent are categorized as being overweight and 9 percent are categorized as being obese. Statistics on the total population show that approximately 55 percent of the Danish adult population is overweight, where 12-13 percent of these are obese (Bendixen, 2003). The difference between the latter statistics and the fraction of people being overweight and obese in our sample is mainly due to the fact that wage-earners tend to be less overweight and obese compared to people out of work and older people.

> INSERT TABLE 1 HERE <

Table 3 presents some simple descriptive statistics on health and health care use in the different weight categories. As mentioned previously, being overweight or obese is linked to a long list of diseases. For that reason we expect people with higher BMI to have worse health, more diseases, and to use more health care than those with lower BMI. The results in Table 3 confirm this when we compare the sample statistics of normal weighted to those of overweight and obese people. Those categorized as obese in 2000 have significantly (at 5 percent) more health care usage than normal weight people, looking at GP visits in 2000 and 2005, hospitalizations in 2000 and 2005, and bed days while hospitalized in 2000 and 2005. Furthermore, obese people more often report having hypertension, diabetes, and heart attacks than normal weight people. Those categorized as being

overweight in 2000 also have significantly more hospitalizations and more bed days in 2000 than those people who fall into the category of normal weight. However, overweight people have significantly (at the 10 percent level) *fewer* GP visits in 2000 than normal weight people.

> INSERT TABLE 2 HERE <

Overall the descriptive statistics in Table 3 show some significant differences in health and the use of health care within different weight categories which has motivated this research. On the one hand, overweight people report *no more* illness (prevalence of hypertension, diabetes, cancer, stroke, and heart attack) than the normal weighted, use *less* medical care (GP visists in 2000), but tend to be hospitalized *more* often than the normal weighted. These results give some support to the theoretical hypotheses mentioned in Section 2, i.e., that obese people have inconsistent preferences and tend to postpone doctor visits unless they are forced to go to the doctor.

On the other hand, Table 3 shows that obese people are more ill and use more medical care. But, not all overweight and obese people are ill and need medical care, a fact which is important to take into account when studying the demand for medical care by these groups. Therefore treating all people similarly may obscure the effect unhealthy bodyweight has on medical care. Therefore, the chosen model for estimation is a finite mixture model where the underlying unobserved heterogeneity (frailty) which splits the population into latent classes is based on the individual's health status.

> INSERT TABLE 3 HERE <

4. Empirical Model
4.1 Empirical strategy

Figure 1 shows the frequency distribution of primary physician visits in the selected sample. The number of GP visits takes on non-negative integer count values within a fixed period, and, has a not insignificant number of zero visits. Figure 1 shows that 18 percent of the sample has zero visits in 2000. Most observations are concentrated around lower values while the right tail is very long with a maximum value of 69. There are only few observations on these higher values.

> INSERT FIGURE 1 HERE <

Two-part models (TPM) have been proposed as a solution in the literature in which the decision to go to a doctor is modeled as distinct from the frequency of visits (Manning et al., 1981; Pohlmeier and Ulrich, 1995). In a TPM, the first part is often modeled as a binary probability model that describes the distinction between users and non-users. The second part is often modeled as an ordinary least squares (OLS) model which describes number of visits conditional on going to the doctor.[7] The theoretical basis for the TPM is the principal-agent model. These models appropriately describe a setting in which the first visit is patient-initiated (part one), while the doctor chooses the intensity of the treatment (part two) (Pohlmeier and Ulrich, 1995). However, many individuals may have several treatment spells starting before and ending after our sample period (1th of January-31th of December 2000). In fact, we might expect that only few individuals have only one treatment spell which both starts and ends within the sample period. Also few outliers, with many doctor's visits, may in a TPM influence the results. For these reasons, the two-part estimation method may not be the optimal choice of estimation strategy here.

In the literature on demand for health care, count models, such as the Poisson or Negative Binomial model (NBM), have been proposed in modeling the demand for GP visits (Teckle and Sutton, 2008). However, modeling all doctors visits (above zero) in the same model raises the problem of heterogeneity, as those who infrequently go to the doctor and those who frequently go to the doctor may vary systematically on unobserved variables, such as their health capital.

More recently, finite mixture models (FMM) have been proposed in the estimation of the demand for health care, also called latent class models (Deb and Trivedi, 1997, Deb and Trivedi, 2002 and d'Uva, Jones and Doorslaer, forthcoming).

One of the advantages of finite mixture models is that one can distinguish between groups with high and low average demand for health care. It is intuitively appealing to treat people with zero or few GP visits a year differently from people with many GP visits a year because the two groups may differ depending on health capital and health behavior. If we hypothesize that there are two kinds of

---

[7]The OLS assumes that the error term is normally distributed. This may lead to a prediction of negative values for the dependent variable which by nature only can take non-negative integer count values. For that reason, OLS does not seem to be appropriate in this setup. When we estimate the TPM in appendix we use the negative binomial model to estimate part two.

users of health care, and that some unobserved heterogeneity/frailty such as the individual's health capital and health demand history divide the data into these two groups, then we can detect the endogeneity exactly by splitting the sample into two components. Different behaviors and factors can then be allowed to have a differential effect on the use of health care for frequent and infrequent users and unobserved factors are also allowed to differ for the two (or more) disaggregated groups. The TPM also splits data into two groups. But the FFM is more flexible than the TPM because we estimate a mixing parameter which forms the components. This means that in the FMM, one can estimate people with zero, one, two and three GP visits a year together and not, as in the TPM where those with few GP visits are estimated together with those with many, e.g., 69, GP visits a year. With enough components the FMM actually becomes semi-parametric. However, experience suggests that usually few latent classes are needed to approximate the density well (Deb and Trivedi, 2002). When the demand for medical care for obese people is estimated via FMM, the unobserved heterogeneity that exists among obese people with respect to health problems is revealed. Therefore, we choose to estimate a FMM and test it against the TPM and NBM.

In the estimation of the FMM, a list of variables is included which measure life style behavior (body weight, smoker status, and smoking intensity and duration). Though some endogeneity problems are taken care of, when using FMM, by splitting the sample into groups by latent health status, the lifestyle variables, i.e., indicators of obesity and overweight, may be endogenous within the two groups. One can easily think of unobserved factors reflecting time or risk preferences which might correlate with both health care demand and lifestyle factors within the two groups. To control for this kind of endogeneity, we have omitted the lifestyle variables one by one in the FMM estimation. Leaving the lifestyle variables out in the estimation gives only small and insignificant changes to the parameter estimates (these results are available upon request). These results indicate that only minor endogeneity problems exist within each group.

In the FMM, as well as in the TPM and NBM, the probability distribution of the data needs to be specified. In this paper the negative binomial distribution is chosen for the mixture component densities. Several other probability distributions (e.g., Normal and Poisson) have been tested, but the negative binomial works best and seems useful also because it allows for correction of over-dispersion and loosens the Poisson restriction by allowing the arrival rate to vary systematically.

The FMM is tested against both the NBM and the TPM in section 5.1. But we will focus on the finite mixture model to estimate the demand for health care because the finite mixture model is a generic model which takes into account the heterogeneity in the data.

4.2 Estimation equation

To estimate the FMM, a sample of the population, N, is used where N is sufficiently large. For each individual in the population, information exists on a number of characteristics including visits to the general practitioner (GP), which is denoted y_1, y_2,...., y_N. The population can now be divided into j subpopulations, j=1, 2...C, with weight π_j so that $\sum_j \pi_j = 1$ and $0 \leq \pi_j \leq 1$. Here we have j=2 which means that we split data into 2 subpopulations--one with infrequent visits to the GP and one with frequent visits to the GP.[8]

For each subpopulation (also called component) a probability density function is composed, which gives the probability of observing y_i in the subpopulation j. Let f_j denote the density function of the subpopulation j and $f_j(y_i)$ denote the probability that observation i is seen in subpopulation j. Then the density is given by:

$$(1) \qquad f(y_i|\Theta) = \sum_{j=1}^{C-1} \pi_j f_j(y_i|\theta_j) + \pi_C f_C(y_i|\theta_C)$$

The mixing parameters, π_j are estimated along with all the other parameters.

The mean of a finite mixture density is given by:

$$(2) \qquad E(y_i|x_i) = \sum_{j=1}^{C} \pi_j \lambda_j$$

Where $\lambda_j = E(y_i|j)$

The posterior probability that observation y_i belongs to the component j is given by:

$$(3) \qquad \Pr(y_i \in subgroup\ j) = \frac{\pi_j f_j(y_i|x'_{i'}, \theta_j)}{\sum_{j=1}^{C} \pi_j f_j(y_i|x'_{i'}, \theta_j)}$$

---

[8] One can easily think of dividing the sample into more than two groups. We have estimated a model with 3 components (subgroups). Specifying 3 components does not significantly improve the model and is much more time-consuming to estimate. Furthermore, and most importantly, it seems intuitively more appealing to divide the sample into infrequent and frequent users which also is the approach generally adopted in the literature (see Deb and Trivedi, 2002).

The component distribution in a finite mixture negative binominal model is specified as:

$$(4) \qquad f(y_i|\theta_1,\theta_2,..\theta_C;\pi_1,\pi_2,..\pi_C) = \sum_{j=1}^{C} \pi_j \frac{\Gamma(y_i+\psi_{j,i})}{\Gamma(\psi_{j,i})\Gamma(y_i+1)} \left(\frac{\psi_{j,i}}{\lambda_{j,i}+\psi_{j,i}}\right)^{\psi_{c,i}} \left(\frac{\lambda_{j,i}}{\lambda_{j,i}+\psi_{j,i}}\right)^{y_i}$$

Where Γ is the gamma function, $\lambda_{j,i} = \exp(x'_i \beta_j)$ and $\psi_{j,i} = (1/\alpha)\lambda_{j,i}^k$. The parameter α is an over-dispersion parameter and k is an arbitrary constant fixed at 1 and 0 in the negative binomial model 1 and 2, respectively.

If the data are assumed to be iid, the joint likelihood of the samples can be defined as the product $\prod_j f(y_i \mid \theta_1, \theta_2,.. \theta_C, \pi_1, \pi_2,... \pi_C)$. The log likelihood function is then:

$$(5) \qquad L(\Theta) = \sum_{i=1}^{N} \log(f(yi \mid \Theta)$$

The model is estimated by maximum likelihood using the Newton-Raphson maximization algoritm.

4.3 Model selection

We test the negative binomial model (NBM), the two part model (TPM), and the finite mixture model (FMM) against each other to see which one performs better. The Akaike Information Criterion (AIC) and Baysian Information Criterion (BIC) are used to compare the three models. The definitions of AIC and BIC are AIC=- ln L +2K and BIC= -2 ln L + K ln(N) where K is the number of parameters in the model, and N is the number of observations. We pick the models with the smallest value of AIC and BIC. Both criteria represent a trade-off between 'fit', as measured by the log likelihood, and 'parsimony', as measured by the free parameters. In terms of parsimony, the FMM is penalized, as more parameters are estimated in the FMM than in the TPM and NBM.

5. Results

5.1 Estimation models

To estimate the demand for medical care a finite mixture model (FMM) is estimated with both negative binominal-1 (NB1) and negative binominal-2 (NB2) based densities. The FMM is compared to a negative binomial model (NBM) and a two part model (TPM). We test these models against each other, allowing the densities to be successively based on NB1 and NB2. To be able to compare the three models, we estimate part two of the NBM and TPM (which estimates the number of doctor visits conditional on going to the doctor) with NB1 and NB2 based densities, as this minimizes

variation from other sources. Part one in the TPM (which estimates the decision of going to the doctor or not) did not fulfil the convergence criterion using either NB1 or NB2 densities. Therefore, part one in the TPM is estimated with a probit, as the choice of model is typically not important where binary models are concerned. Deb and Trivedi (2002) estimate part one with both logit and NB1 and obtain almost identical results.

5.2 Model selection

Table 4 present the selection criteria log L, AIC and BIC for the NBM, TPM and FMM based on NB1 and NB2 densities for the estimation of medical care. When comparing the models in Table 4, the FMM, based both on NB1 and NB2, is the preferred model compared to NBM and TPM, as this model has the lowest AIC and BIC.

For the NBM and FMM model, NB1 densities fit the data better, i.e., both log L, AIC and BIC has a lower value when based on NB1 densities than on NB2 densities. The TPM based on NB2 is favoured over the TPM based on NB1, but the TPM is the least preferred model in Table 4. Overall, the results in Table 4 show that the FMM based on NB1 density is the preferred model.

> INSERT TABLE 4 HERE <

5.3 Results and test of FMM

In Table 5, the estimate for the mixing parameter, which denotes the weight given to each subpopulation, is labelled π. This parameter shows that the probability of being an infrequent user of medical care (in component 1) is 0.60. This means that 40 percent of all people in the sample are considered as frequent users (their probability of being a frequent user is above 50). In Table 6 the minimum value of primary physician visits among infrequent users is 1.2, maximum value is 14.5, and mean value is 3.3. The corresponding numbers for frequent users are 1.7, 26.2, and 6.6. These results show that the distributions for infrequent and frequent users differ significantly and support our assumption of heterogeneity in the sample.

5.3.1 Overweight, obesity and use of primary physician care

Table 5 presents the results from the FMM estimating the number of GP visits in 2000. A central finding is that conditional inference matters: being overweight or obese does not increase the use of

primary physician care among the infrequent users, whereas being overweight and obese increases the use of primary physician care among frequent users. However, the magnitude of the parameter estimate on overweight and obesity among frequent users is relatively small compared to, e.g., the parameter estimate on income. Overweight and obese people, in the group of frequent users, have on average 0.09 and 0.17 more GP visits than normal weight people. The parameter estimate on wage shows that a 10 percent increase in wages reduces the number of GP visits by 1.39. However, the estimates on overweight and obesity are the same as the effect of an increase in age on doctor visits of 10 and 16 years, respectively.[9]

In section 2, both economic and non-economic theories provided explanations as to why people with higher bodyweight would not demand more primary physician care when they are relatively healthy (when they do not have an acute disease) than normal weight people although they have a much higher probability of getting several lifestyle diseases. The results in Table 5, to some extent, support these theories, as infrequent users in the overweight and obese weight categories do not use more physician care than normal weight people, despite the fact that they face a higher risk of developing a long list of lifestyle diseases.[10]

> INSERT TABLE 5 HERE <

As the primary physician acts as the gatekeeper to specialist consultations in the Danish health care system, one might worry that the people categorized as infrequent (and overweight or obese) are categorized as such because while they make fewer visits to the GP, they may make more specialist visits. Table A1 shows mean visits to specialists (which include anesthesiologist, radiologist, dermato-venereologist, physiatrist, gynaecologist, surgeon, neurosurgeon, orthopaedic surgeon, pathologist, plastic surgeon, and psychiatrist) and other specialists (physiotherapist, chiropractor, podiatrist, psychologist/therapist, and medical treatments where the doctor visits the patient) and

---

[9] We have tested different BMI cut-offs to measure overweight and obesity (e.g., BMI cut-off at 27 and 32). These different BMI cut-offs did not significantly change the sign and magnitude for the group of overweight people. However, when the BMI cut-off is increased to 32 (or 35) to measure obese people only 4 (or 2) percent of the sample are in the obese category, and being obese does not increase the use of physician care among frequent users, but does so among infrequent users. These results are available upon request.

[10] To test whether individuals in the overweight and obese weight category are in fact healthy, we have tried to include in interaction of self-rated good health and weight category. These results, however, are somewhat redundant in the FMM as the two components already measure an interaction between latent health and obesity category. Still, the only significant interaction in the FMM show that given a good self-rated health, frequent users in the overweight category have 0.25 *less* GP visits a year than normal weight people. In the Negbin and Two-part model the interaction effect is not significant.

tests for differences between frequent and infrequent users. Among infrequent users both overweight, obese, and all have significant fewer specialist visits than frequent users. These results suggest that overweight and obese people do not substitute specialist care for GP care.

However, many other kinds of therapists also exist outside the general health service system. In particular, several alternative treatment methods are available offering diet and exercise instructions (e.g., weight watchers). Table A1 shows that among frequent users, obese and overweight people use significantly more alternative treatments to control diet and receive exercise instruction than among obese and overweight people among infrequent users. Furthermore, among both frequent and infrequent users, normal and overweight people have exactly the same tendency to use alternative treatment methods offering diet and exercise instruction.

5.3.2 Other background factors and use of primary physician care

The estimation of the FMM in Table 5 shows that the parameter estimates on male, wages and good health have the same sign and significance for infrequent and frequent users. This indicates robustness of the parameter estimates of these variables for the whole sample. Men make significantly fewer GP visits than women[11], people in self-reported good health have significantly fewer GP visits than those with poor health, and high wage earners make fewer GP visits than low wage earners. This last result can be explained by greater time costs of doctors' visits of high-wage earners, i.e., the foregone earnings associated with visiting a GP as money costs of medical care are not significant in a system where most physician (including specialist) visits and hospitalizations are universal and free of charge. High wage-earners might also simply be in better health than low wage-earners.

Self-reported good health is related to overweight and obesity, as a high BMI is a risk factor for diseases. The total effect of overweight and obesity could be obtained by excluding variables measuring good health. However, we retain self-reported good health as it explains a significant proportion of the demand for medical care, and when this variable is left out, the parameter estimates on overweight and obesity include the impact of health on the demand for primary physician care.

---

[11]Previous studies have also found that women's health care utilization is significantly greater than men's whether measured by expenditures or by quantities (Sindelar, 1982).

> INSERT TABLE 6 HERE <

Among infrequent users, younger people make more GP visits whereas among frequent users, older people make significantly more GP visits. Many diseases are associated with aging, so we would expect older people to use more medical care. For that reason, among infrequent users, older people are less likely to use GP services because they are in relatively good health.

Small children are more often sick, and we therefore assume that people with small children have more health knowledge because they use the health care system more often than people without small children. Among infrequent users, people with small children use more health care.

Being an immigrant in Denmark or having an emotionally stressful job induces infrequent users to use more medical care. This means that, even if people are in relatively good health, these two characteristics imply a need for more medical care.

Among frequent users, married or cohabiting people use less medical care than single people. It is difficult to interpret this result as the causal effect of marriage, as this could be measuring the effects of different factors. For example, that healthy people are selected into marriage, or that the state of marriage confers health protection because it allows for greater economic resources and may have an enforcing effect on healthy behaviours.

Lifestyle behaviours have a different impact on medical care among infrequent versus frequent users. Among infrequent users, being a current smoker reduces the number of visits at the GP. One would expect smokers to use more medical care as several diseases are related to smoking. However, this result might be explained by the fact that young people are more likely to be in the group of infrequent users (see table A1), and the complications related to smoking have therefore not been discovered yet. Still, people who have smoked 6 to 10 years go significantly more to the GP than people who have smoked only 1 to 5 years.

5.3.3 Testing the FMM

18

In Table 7, the estimated parameters across frequent and infrequent groups are tested for equality by way of a Wald test. One advantage of the FMM is that by splitting the sample into two groups, we capture the heterogeneity in the data which originates from individuals' unobserved health capital and health behaviour. One would therefore expect the parameter estimates of the included health variables to differ across the two components.

Table 7 shows that the FMM seems to perform well, as significant differences are found in the parameter estimates for the two groups. The parameter estimates for frequent and infrequent users differ significantly at the 5 percent level for the variables measuring age, marital status, self-rated good health, and smoking. Consequently, we interpret the latent classes in the FMM as differences in health status and health behaviour. As noted earlier, also marital status could represent a health effect.

> INSERT TABLE 7 HERE <

Table A2 in the appendix presents the results on the demand for medical care estimated by the NBM and the TPM based on NB1 densities. The parameter estimates on gender, wage and good health have the same signs and significances as in the FMM which confirm the robustness of these results. Many of the included variables are significant in the NBM and the TPM. Because people in the tails of the distribution with many GP visits also are those with, e.g., bad health and health behaviours, these variables will have a significantly larger impact on the model when we estimate all users in the same model, as in a NBM or TPM, compared to the FMM. Consequently, the NBM and the TPM give very limited information on infrequent (or relatively healthy) users.

To further test our FMM for evidence on heterogeneity dependent on latent health status, we follow Deb and Trivedi (2002) and estimate the logarithm of the estimated posterior probability that individual i belongs to the group of infrequent users (see equation 3 in section 4.1) on the same list of control variables previously used in the FMM plus additional indicators of the individual's health and health behavior. The additional variable measuring health is whether a doctor has diagnosed interview person (IP) with a disease and the additional variables measuring health behavior are separate indicators equaling respectively 1 if IP never eats vegetables and fruit, and 1 if IP never

exercises. The additional variables serve as an external check of the characteristics of the latent class, which we have claimed reflect health and health behavior.

Table 8 shows the results from the posterior regression. These results show that higher posterior probabilities of belonging to the group of infrequent users are positively associated with being younger, married, in good self-reported health, doing some kind of exercise, and not having a diagnosis reported by a doctor. If we believe that being married is a measure of a good health status, then the association between the posterior probability and the list of variables just mentioned support our hypothesis that individuals in the two latent classes may be classified as healthy and unhealthy individuals.

It is not clear how to interpret the smoking variables as the indicator for missing information on smoking duration is significant and positively associated with higher posterior probabilities of belonging to the group of infrequent users. This means that the result that non-smokers and individuals who have smoked 6-10 years are more often infrequent users than people who have smoked 1-5 years, only is valid for the selected sample of people answering the question on smoking duration.

A higher posterior probability of belonging to the group of infrequent users is also associated with higher wages. This association could be reflective of higher time costs of going to the doctor for high wage earners and therefore that their use of medical care is lower. However, the positive relation between a high posterior probability of belonging to the group of infrequent users and wages could also indicate that high wage earners are in relatively better health, which again supports our hypothesis that the frailty which splits the two latent classes is based on health and health behavior.

> INSERT TABLE 8 HERE <

6. Discussion

In section 2, a number of potential explanations were suggested for why we might observe that increasing bodyweight does not lead to increased demand for primary physician care among infrequent users, as both the empirical and the descriptive evidence have shown. Higher bodyweight

may correlate with higher discount rates. Similarly, time inconsistent preferences may affect both body weight and an activity such as visiting the doctor. It is difficult to find convincing proxy measures for the individual's discount rate and inconsistent time preferences, and we do not access measures of functional limitation either, which could have shed light on whether difficulty walking or climbing stairs may act as a barrier to preventive health care checkups. Neither can we make a direct test whether discriminatory attitudes encountered during visits or even the perception of discrimination could discourage the obese from utilizing preventive medical care.

However, from the medical literature we know that not all obese individuals are genetically disposed to the metabolic syndrome, and not all obese people are sick. Through the mixture model we can separate the latent classes determined by health status and health behaviour.

We find that conditional inference matters, i.e., obese and overweight individuals do not demand more primary physician care than normal weight individuals when they are healthy (infrequent users) but do demand more health care than normal weight individuals when they are unhealthy (frequent users). These results support the hypotheses which were described in Section 2. Grossman's model predicted an increase in the demand for medical care due to obesity, and we find that when overweight and obesity is complicated by sickness, the use of medical care does increase (for frequent users). The hypothesis on time inconsistent preferences predicted a lower demand for medical care among overweight and obese people. The results from the FMM showed that when overweight and obese people are relatively healthy (infrequent users) the use of medical care is not different from normal weight people. One interpretation of this result may be that obese and overweight individuals do not appear to utilize *preventive* medical care. Rather, they tend to demand medical care only when illnesses are diagnosed.

The fact that it is preventive care that differs between infrequent and frequent users is supported by descriptive statistics in Table A1. Table A1 shows that frequent users, within all weight categories, use more preventive care than infrequent users. Frequent users have also their blood pressure and cholesterol tested more often than infrequent users. Furthermore, Table A1 shows that there is no difference in preventive care among the different weight categories among infrequent users. This means that among infrequent users people in different BMI categories have the same amount of

21

preventive care. Notice, that there are significant numbers of obese and overweight people among infrequent users (33 and 8 percent, respectively).

Does the lack of preventive care on the part of the obese impose a cost on taxpayers? In Table 9, we present an out-of-sample validation based on register information on health care usage in 2000 and 2005 according to the individual's weight category and intensity of use in 2000, the latter being predicted from the latent classes identified in the model. Individuals can be followed over time because the NHI survey sample is linked to administrative registers up to 2005. In 2005, among infrequent users, which can be considered the target group for preventive medical care, neither obese nor overweight individuals make more GP visits than normal weight individuals. In fact, overweight individuals make significantly *fewer* GP visits than normal weight individuals at the 5% level.

Five years later, in 2005, it can be seen that individuals who were obese and infrequent users in 2000 make at least one GP visit per year more, have twice as high a rate of hospitalization and four times as many bed days than normal weight individuals. Furthermore, both obese and overweight infrequent users are also hospitalized significantly more often with cardiovascular disease five years later compared to their normal weight counterparts. There are no significant differences, however, in cancer rates[12] most likely because there are simply too few observations in these states. Among frequent users of GP care in 2000, it can be seen from the second panel in Table 9 that while the obese do not make more primary physician visits than normal weight individuals in 2000, they make about 2 GP visits more a year than normal weight individuals five years later.

A substantial *increase* in health care usage of the obese (and to a lesser extent of overweight individuals) occurs between 2000 and 2005 among infrequent users. Infrequent users who are obese make nearly twice as many doctors' visits, and have 5 times as many hospitalizations, 9 times as many bed days and 3-4 times as many hospitalizations for cardiovascular disease in 2005 than in 2000. Whereas among frequent users, health care usage decreases slightly or stays the same within weight categories over this period. This difference between frequent and infrequent user cannot only be due to a pure age effect, i.e., infrequent users are just frequent users five years before. Although the sample of those who are categorized as being infrequent users on average are

---

[12]There were too few observations in the case of diabetes and hypertension to be able to make these comparisons.

significant younger than those who are frequent users, the age difference is only 3.8 years (see table A1). Thus, the out-of-sample validation indicates that the fact that obese and overweight individuals do not demand preventive health care may imply a higher rate of health care utilization and greater hospitalization costs in the future, which in the end imposes a cost on taxpayers in the Danish health care system which offers universal and costless access to medical care.

> INSERT TABLE 9 HERE <

7. Conclusion

In this paper, we investigate the relationship between primary physician care, overweight and obesity in 2000 and the predicted out-of-sample impact of these observed patterns of health care use and associated indirect costs five years later.

Our choice of econometric model has implications for the interpretation of the association between obesity and medical care. We use a finite mixture variant of the latent class model to investigate the relationship between obesity and demand for medical care (Deb and Trivedi, 1997; Deb and Trivedi, 2002). By characterizing the demand for health care via a mixture distribution, we split the population into frequent and non-frequent users of primary physician services according to the individual's latent health status, health risks, and attitudes. The finite mixture model is used to learn whether obese individuals engage in more preventive medical care—e.g., by demanding medical care when they are healthy–or whether they tend to demand more medical care only when unhealthy, i.e., when illnesses are diagnosed.

We find that obese people tend not to use more GP services when they are relatively healthy (infrequent users) than people in the normal weight group. While we cannot directly identify the underlying causes for the lack of correlation between body weight and GP care, we do observe that among infrequent users of primary physician care in 2000, the use of medical care five years later is significantly larger for obese people than normal weight people. These results suggest that people with higher body weight have inconsistent time preferences as they do not use more preventive care in the present time even though they face more health risks in the future. Furthermore, these results

23

indicate that more preventive care for selected people within this group would reduce hospitalization costs.

That relatively healthy (infrequent users) overweight and obese people use no more medical care than normal weight people might also be explained by the fact that young people are more likely to be more healthy (and in the group of infrequent users). Because of their age, the obese people in the infrequent group might be undiagnosed 'silent carriers' of obesity related diseases. As Denmark has one of the highest rates of obesity in the EU countries in the 15-29 age interval, targeting preventive care at this group might reduce future health care costs related to overweight and obesity.

References

Bendixen, H. 2003. Changes in prevalence of overweight, obesity and slimming behaviour in Danish adults with emphasis on dietary fat quality (ph.d.-dissertation). Faculty of Life Science, Copenhagen

Bhattacharya, J., Sood, N. 2005. Health Insurance and the Obesity Externality. NBER Working Paper W11529

Borghans, L., B.H.H. Golsteyn. 2005. Time Discounting and the Body Mass Index. IZA Discussion Paper Series 1597

Cawley, J. 2006. Reefer Madness, Frank the Tank or Pretty Woman: To What Extent do Addictive Behaviors Respond to Íncentives? Paper presented for the conference "Incentives and Choices in Health Care: Contributions and Limitations of Economics" August 2006

Cropper, M.L. 1981. Measuring the Benefits from Reduced Mortality. *American Economic Review*, 71, 235-240

Cutler, D., E. Glaeser. 2005. What Explains Differences in Smoking, Drinking and other Health-Related Behaviors? *American Economic Review*, 92(2): 238-242

Deb, P., P. K. Trivedi. 1997. Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics*, 12: 313-336

Deb, P., P. K. Trivedi. 2002. The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics* 21: 601-625

d'Uva, T. B., A. Jones, E. Doorslaer. forthcoming. Measurement of horizontal inequality in health care utilisation using European panel data *Journal of Health Economics* forthcoming 2009

Ekholm, O. et al. 2006. Sundhed og sygelighed I Danmark 2005 (The Danish Health and Morbidity Survey 2005). National institute of Public Health , Copenhagen, December 2006

Grossman, M. 1972a. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80, 223-255

Grossman, M. 1972b. The Demand for Health-A Theoretical and Empirical Investigation. New York, Columbia University Press

Grossman, M., Rashad, I. 2004. The Economics of Obesity, *Public Interest*, 156: 104-112

Hoppe R., Ogden J. 1997 Practice nurses' beliefs about obesity and weight related interventions in primary care. *Int J Obes Relat Metab Disord*. 21(2):1416

Ismaa, B. 2003. A major health hazard: The metabolic syndrome. *Life Science*, Volume 73, Number 19, 26: 2395-2411(17)

Komlos, J., P. Smith, B. Bogin. 2004. Obesity and the Rate of Time preferences: Is there a Connection? *Journal of Biosocial Science*, 36(2): 209-219

Lakdawalla, D., Philipson, T.J. 2006. Economics of Obesity. In: Jones, A. (Ed) Elgar Companion to Health Economics, 72-82

Manning, W.G. et al. (1981). "A Two-Part Model of the Demand for Medical Care: Preliminary Results from the Health Insurance Study." In Health, Economics and Health Economics, ed. Jacques van der Gaag and Morris Perlman, 103-23, Amsterdam: North Holland.

Masbad, S. and Astrup, A. 2004. Fedme, metabolisk syndrom og hjerte-kar-sygdomme (Obesity, the metabolic syndrome, and cadiovascular diseases), *Ugeskrift for læger* 166/17.

Ministry of Health and Prevention. 2007. De samfundsøkonomiske konsekvenser af svær overvægt. (The Economic Consequences of Obesity). Report Ministry of Health and Prevention May 2007

NIH, NHLBI Obesity Education Initiative. 1998. "Clinical Guidelines on the Identification, Evaluation, and Treatment of Overweight and Obesity in Adults." Available online http://www.nhlbi.nih.gov/guidelines/obesity/ob_gdlns.pdf

Nocera, S., Zweifel, P. 1997. The Demand for Health: An Empirical Test of the Grossman Model using Panel Data. In: Zweifel, P. (Ed.) Advances in Health Economics, Dordtrecht: Kluwer, 35-49

OECD. 2001. Employment Outlook. June 2001. OECD, Paris

Pohlmeier, W. and Ulrich, V. (1995). "An Econometric Model of the Two-part Decisionmaking Prcoess in the Demand for Health Care", *The Journal of Human Resources*, Vol. 30, Nol 2 (Spring, 1995), pp. 339-361.

Price J.H., Desmond S.M., Krol R.A., Snyder F.F., O'Connell J.K. 1987. Family practice physicians' beliefs, attitudes, and practices regarding obesity. *Am J Prev Med*, 3(6):33945

Puhl, R., K. D. Brownell. 2001. Bias, Discrimination and Obesity. *Obesity Research*, 9, 12: 788-805

Sanz-de-Galdeano, A. 2005. The Obesity Epidemic in Europe, IZA DP, 1814

Sculati O., Bettoncelli G., Brignoli O., Corgatelli P., Ponti D., Rumi A., Zucchi A. 2006. Efficient prevention of overweight and obesity in the experience of family practitioners and "nutrition units" of the public health system in Lombardy. *Ann Ig*. Jan-Feb;18(1):41-8

Sindelar, J. L. 1982. Differential Use of Medical Care by Sex. *Journal of Political Economy*, 90,5:1003-1017

Statistics Denmark. 2001. Sociale forhold, sundhed og retsvæsen; sygesikring 2000. Statistiske efterretninger 2001:16

Smith, P. K., Bogin, B., Bishai, D. 2005. Are Time Preference and Body Mass Index Associated? Evidence from the National Longitudinal Survey of Youth. *Economics and Human Biology*, 3: 259 - 270

Svendsen, L., Heitman, B.L., Mikkelsen, K.L., Raben, A., Ryttig, K.R., Sørensen, T.I.A., Breum, L., Tauber-Lassen, E., Hansen, G.L. , Ovesen, L.F. , Richelsen, B., Søndergaard, K., Toubro, S.. 2001. Fedme i Danmark.(A Report from "Task Force on Obesity"). Danish Association for the Study of Obesity. Nr 8.

Teckle, P, M. Sutton. 2008. How do the determinants of demand for GP visits respond to higher supply? An analysis of grouped counts. Unpublished. Presented at SSES Annual Meeting. Health Economics and Econometrics. Lausanne, March 27-28 2008

Van Baal, P.H.M., Polder, J.J., de Wit, G.A., Hoogeveen, R.T., Feenstra, T.L. 2008. Lifetime Medical Cost of Obesity: Prevention No Cure for Increasing Health Expenditure. *PLoS Medicine*, 5,2

Wagstaff, A. 1986. The Demand for Health: Some New Empirical Evidence. *Journal of Health Economics*, 5(3): 195-233.

Wolf A.M., Colditz, G.A. 1998. Current estimates of the economic cost of obesity in the United States. *Obesity Research*, 6(2):97–106

Table 1: Detailed Variable Description

Variable	Definition
Health care use	
GP visits 2000	Number of visits to the general practitioner in year 2000. Register data
GP visits 2005	Number of visits to the general practitioner in year 2005. Register data
Specialist	Number of visits to specialists in 2000. Specialists include: anesthesiologist, radiologist, dermato-venereologist, physiatrist, gynaecologist, surgeon, neurosurgeon, orthopaedic surgeon, pathologist, plastic surgeon, and psychiatrist. Register data.
Other specialist	Number of visits to other specialists in 2000. Other specialists include: physiotherapist, chiropractor, podiatrist, psychologist/therapist, and doctors who visit their patients. Register data.
Hospitalizations 2000	Number of hospitalizations with a diagnosis in year 2000. Register data
Hospitalizations 2005	Number of hospitalizations with a diagnosis in year 2005. Register data
Bed days 2000	Number of bed days in year 2000. Register data
Bed days 2005	Number of bed days in year 2005. Register data
Preventive care	IP had a preventive physical examination within the last year
Blood pressure	IP had blood pressure measured at the GP within the last year
Cholesterol	IP had blood cholesterol measured within the last year
Diet and exercise instruction	Whether IP had been treated by therapists outside the general health service system and received instruction regarding diet and exercise
Health	
Diagnosis 2000	Self-reported diagnosis
Doctor diagnosis	Self-reported whether doctor reported diagnosis
Hypertension	Self-reported information on hypertension
Diabetes	Self-reported information on diabetes
Cancer	Self-reported information on cancer
Stroke	Self-reported information on stroke
Heart attack	Self-reported information on heart attack
Good health	IP rates health as very good or good.
Health behavior	
No exercise	During the last year IP did only passive physical activities and no exercise
No vegetables	IP never or very seldom eats vegetable or fruit
Overweight	Overweight (25 <= BMI < 30) in 2000
Obese	Obese (BMI>=30) in 2000
Smoke	Current smoker
more15smo	Smoked more than 15 cigarettes a day (both previous and current smokers)
Smoking duration	
Smoked0yr	Never smoked
Smoked1_5yr	Smoked 1-5 years
Smoked6_10yr	Smoked 6-10 years
Smoked11_yr	Smoked more than 11 years
Smokedmisyr	Smoked, but missing information on number of years

Table 1 continued	
Variable	**Definition**
Other variables	
Contraception	If IP is a woman and has been prescribed use of contraception
Basic	Basic schooling. Register data
Vocational	Vocational training or high school. Register data
Further	Further schooling. Register data
Age	Age at interview. Register data
Children	Number of children in IP's family. Register data
Small child	Dummy if there are children below 10 years in IP's family
Married	Married or cohabiting. Register data
Immigrant	Immigrant. Register data
Logwage	Log of total wage income
Emotional	Q: "How often do emotionally stressing situations arise for you at work?": 1 if IP answers always or often
Physical	IP is often exposed to physically demanding work such as carrying heavy objects and vigorous shaking or vibrations in the body.
County	
Copenhagen	IP lived in Copenhagen county in 2000
Frederiksborg	IP lived in Frederiksborg county in 2000
Roskilde	IP lived in Roskilde county in 2000
Vest Sjaelland	IP lived in West Sjaelland county in 2000
Storstroms amt	IP lived in Storstroms county in 2000
Bornholm	IP lived in Bornholms county in 2000
Fyn	IP lived in Fyns county in 2000
Sønderjylland	IP lived in Southern Jylland county in 2000
Ribe	IP lived in Ribe county in 2000
Vejle	IP lived in Vejle county in 2000
Ringkøbing	IP lived in Ringkobing county in 2000
Aarhus	IP lived in Aarhus county in 2000
Viborg	IP lived in Viborg county in 2000
Njylland	IP lived in Northern Jylland county in 2000

Table 2: Descriptive statistics. Mean and Standard error (SE), N=8,036

	Mean	SE
GP visits 2000	4.622	5.346
Contracep	0.104	0.306
Men	0.518	0.500
Vocational	0.435	0.496
Further	0.301	0.459
Age	42.346	9.806
Children	0.806	1.010
Small child	0.317	0.466
Married	0.781	0.413
Immigrant	0.017	0.128
Logwage	12.348	0.607
Emotional	0.152	0.359
Physical	0.573	0.495
Overweight	0.337	0.473
Obese	0.090	0.287
Goodhealth	0.865	0.342
Frkborg	0.069	0.253
Roskilde	0.072	0.258
Vsjaelld	0.061	0.240
Storst	0.053	0.224
Bornholm	0.034	0.182
Fyn	0.065	0.246
Sjylland	0.059	0.236
Ribe	0.068	0.251
Vejle	0.060	0.238
Ringk	0.063	0.244
Aarhus	0.090	0.287
Viborg	0.070	0.255
Njylland	0.068	0.252
Smoke	0.347	0.476
More15smo	0.464	0.499
Smoked0yr	0.401	0.490
Smoked6_10yr	0.036	0.185
Smoked11_yr	0.539	0.499
Smokedmisyr	0.010	0.101

Notes: Omitted categories are basic education, normal weight, Copenhagen, and smoke1_5yr (smoked 1-5 years)

Table 3: Health and health care use in weight categories normal weight, overweight, and obese in 2000. Mean and standards error (SE)

	Normal weight		Overweight		Obese	
	Mean	SE	Mean	SE	Mean	SE
GP visits 2000	4.633	5.330	4.407[+]	5.268	**5.347**	5.666
GP visits 2005	5.418	6.333	5.550	6.598	**6.986**	7.245
Hospitalizations 2000	0.072	0.343	0.088[+]	0.499	**0.113**	0.518
Hospitalizations 2005	0.109	0.679	0.119	0.621	**0.182**	0.828
Bed days 2000	0.226	1.772	**0.331**	2.226	**0.417**	2.550
Bed days 2005	0.342	2.945	0.421	3.013	**1.007**	7.912
Diagnosis 2000	0.651	0.477	0.658	0.475	**0.751**	0.433
Doctor diagnosis	0.168	0.374	0.188	0.391	**0.222**	0.416
Hypertension	0.048	0.214	0.101	0.301	**0.194**	0.396
Diabetes	0.012	0.108	0.016	0.124	**0.037**	0.189
Cancer	0.022	0.148	0.022	0.147	0.028	0.164
Stroke	0.005	0.071	0.007	0.081	0.008	0.091
Heart attack	0.006	0.075	0.013	0.111	**0.022**	0.147
Obs. (% of all)	4,601 (57)		2,709 (34)		726 (9)	

Note: Mean value in bold shows that overweight (obese) differ significantly at the 5 percent level from normal weight
[+]: overweight (obese) differ significantly at the 10 percent level from normal weight

Table 4: Test of different models of the demand for health care (GP visits). Log likelihood, AIC, and BIC, N=8,036

	K	Log L	AIC	BIC
NBM-1	36	-20,451	40,974	41,043
NBM-2	36	-20,537	41,146	41,215
TPM-1	71	-21,116	42,374	42,509
TPM-2	71	-21,023	42,188	42,324
FMM-1	73	-20,371[a]	40,888[a]	41,027[a]
FMM-2	73	-20,400	40,946	41,085

Note: NBM-i=negative binomial. TPM-i=two part model. FMM-i= Finite mixture model, where i=1,2 is NB1 and NB2 based densities. K= number of parameters estimated in the model
Log L= log likelihood, AIC=-2log(L) + 2K, BIC=-2log(L) + K log(N)
[a] The preferred model

Table 5: Results of finite mixture model. Men and women. Parameter estimates. Robust standard error (robust s.e.)

	Component (infrequent users)		Component 2 (frequent users)	
	Estimate	Robust s.e.	Estimate	Robust s.e.
Contracep	0.316***	0.068	0.177	0.089
Men	-0.649***	0.055	-0.383***	0.061
Vocational	0.030	0.050	-0.055	0.055
Further	-0.018	0.063	-0.103	0.067
Age	-0.008**	0.004	0.011***	0.004
Children	-0.057*	0.031	-0.001	0.038
Small child	0.237***	0.070	0.044	0.089
Married	0.082	0.053	-0.133**	0.061
Immigrant	0.261**	0.131	0.009	0.176
Logwage	-0.077**	0.031	-0.139***	0.034
Emotional	0.198***	0.069	0.005	0.072
Physical	-0.025	0.042	0.032	0.048
Overweight	0.006	0.048	0.090*	0.053
Obese	0.074	0.074	0.166**	0.075
Goodhealth	-0.454***	0.061	-0.677***	0.053
Frkborg	0.138	0.093	0.024	0.115
Roskilde	0.156	0.095	0.101	0.111
Vsjaelld	0.049	0.092	0.051	0.105
Storst	-0.068	0.099	0.073	0.102
Bornholm	-0.089	0.119	-0.017	0.125
Fyn	0.006	0.097	-0.109	0.122
Sjylland	0.009	0.100	0.094	0.106
Ribe	0.064	0.085	0.110	0.100
Vejle	-0.042	0.096	0.109	0.104
Ringk	0.139	0.095	-0.008	0.115
Aarhus	-0.010	0.081	-0.033	0.099
Viborg	0.122	0.115	-0.242*	0.134
Njylland	0.029	0.096	-0.130	0.117
Smoke	-0.122**	0.054	0.030	0.061
More15smo	-0.028	0.063	0.072	0.075
Smoked0yr	0.202	0.170	-0.156	0.162
Smoked6_10yr	0.413**	0.201	-0.494*	0.258
Smoked11_yr	0.242	0.159	-0.004	0.160
Smokedmisyr	0.517	0.325	-0.187	0.359
Constant	2.741***	0.410	3.908***	0.443
Π	0,603 (0,099)			
Log L	-20,371			

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: Omitted categories are basic education, normal weight, Copenhagen, and smoke1_5yr (smoked 1-5 years).

Table 6: Summary statistics of fitted values, mean, standard error (SE), Minimum and maximum value, N=8,036

	Mean	SE	Minimum	Maximum
Sample	4.622	5.346	0	69
Overall in FMM	4.599	1.995	1.957	16.556
Component 1 (infrequent users)	3.306	1.670	1.206	14.470
Component 2 (frequent users)	6.565	3.191	1.769	26.244

Table 7: Test for equality of parameters for component 1 (infrequent users) and component 2 (frequent users), selected parameters tested together.

Parameters	Chi squared	p-value
Men	$\chi^2(1) = 9.12$	0.0025
Age	$\chi^2(1) = 13.57$	0.0002
Education	$\chi^2(2) = 1.02$	0.6015
Children	$\chi^2(1) = 0.85$	0.356
Small child	$\chi^2(1) = 1.93$	0.1652
Married	$\chi^2(1) = 5.78$	0.0162
Immigrant	$\chi^2(1) = 0.94$	0.3328
Logwage	$\chi^2(1) = 1.45$	0.2283
Emotional	$\chi^2(1) = 2.42$	0.1197
Physical	$\chi^2(1) = 0.54$	0.4637
Goodhealth	$\chi^2(1) = 6.17$	0.013
Bodyweight variables	$\chi^2(2) = 1.09$	0.5793
Smoking variables	$\chi^2(6) = 20.21$	0.0025
County	$\chi^2(13) = 9.2$	0.7577

Table 8: Posterior regression, OLS estimates, standard error (SE)

	Estimate	SE
Contracep	0.012	0.028
Men	-0.025	0.018
Vocational	0.023	0.019
Further	0.009	0.022
Age	-0.006***	0.001
Children	-0.016	0.012
Small child	0.033	0.026
Married	0.062***	0.020
Immigrant	0.074	0.059
Logwage	0.064***	0.014
Emotional	0.034	0.021
Physical	-0.007	0.016
Overweight	-0.017	0.017
Obese	-0.031	0.027
Goodhealth	0.177***	0.023
Frkborg	0.053	0.034
Roskilde	0.006	0.034
Vsjaelld	-0.001	0.036
Storst	-0.074*	0.038
Bornholm	-0.003	0.045
Fyn	0.050	0.035
Sjylland	-0.092**	0.036
Ribe	-0.020	0.035
Vejle	-0.043	0.036
Ringk	0.029	0.035
Aarhus	0.021	0.031
Viborg	0.071**	0.034
Njylland	0.034	0.034
Smoke	-0.028	0.021
More15smo	-0.013	0.024
Smoked0yr	0.130**	0.065
Smoked6_10yr	0.190**	0.074
Smoked11_yr	0.096	0.065
Smokedmisyr	0.256***	0.098
No exercise	-0.059**	0.024
No vegetables	0.002	0.019
Doctor diagnosis	-0.249***	0.020
Constant	-1.388***	0.176
Adj. R^2	0,05	
N	8,036	

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: Omitted categories are basic education, normal weight, Copenhagen, and smoke1_5yr (smoked 1-5 years)

Table 9: Health and health care use in 2000 and 2005 for frequent and in-frequent users in weight categories normal weight, overweight, and obese in 2000. Mean and standard deviation (SD)

	Infrequent users						Frequent users					
	Normal weight		Overweight		Obese		Normal weight		Overweight		Obese	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
GP visits 2000	3.008	2.908	**2.472**	2.594	2.905	2.875	11.334	7.412	10.861	6.614	11.683	6.195
GP visits 2005	4.443	4.871	4.423	5.076	**5.391**	6.050	9.414	9.356	9.328	9.200	**11.141**	8.384
Hospitalizations 2000	0.039	0.217	0.037	0.242	0.032	0.207	0.204	0.623	0.258	0.921	**0.322**	0.892
Hospitalizations 2005	0.077	0.459	0.095	0.575	**0.159**	0.762	0.239	1.212	0.196	0.749	0.242	0.978
Bed days 2000	0.110	0.855	0.112	1.076	0.088	0.729	0.704	3.578	1.062[+]	4.115	1.272[+]	4.589
Bed days 2005	0.240	2.206	0.303	2.487	**0.810**	5.987	0.757	4.910	0.817	4.313	1.520	11.511
Hospitalized with cancer 2000	0.005	0.071	0.002	0.049	0.006	0.076	0.017	0.128	0.010	0.098	0.005	0.070
Hospitalized with cancer 2005	0.004	0.064	0.003	0.054	0.006	0.076	0.007	0.082	0.008	0.089	0.005	0.070
Hospitalized with cardiovascular disease 2000	0.002	0.049	0.002	0.044	0.004	0.062	0.017	0.128	0.024	0.153	0.025	0.156
Hospitalized with cardiovascular disease 2005	0.005	0.068	**0.008**	0.090	**0.011**	0.106	0.010	0.100	0.018	0.132	0.020	0.140
N	3,703		2,084		524		898		625		202	

Note: Mean value in bold shows that overweight or obese differ significantly from normal weight at the 5 percent level by infrequent or frequent user.

Figure 1: Frequency distribution of GP visits
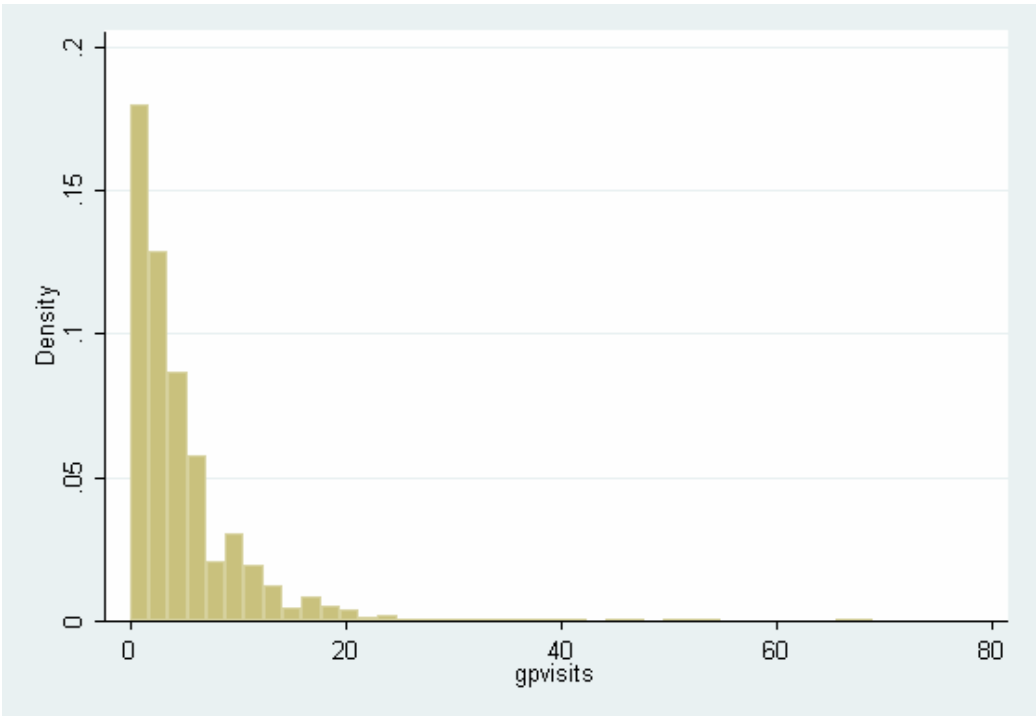
Appendix

Table A1: For selected variables, mean, standard deviation (SD), and test of differences between frequent and infrequent users.

	Infrequent users								Frequent users							
	Normal weight		Overweight		Obese		All		Normal weight		Overweight		Obese		all	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	**40.43**	9.66	**43.23**	9.67	**42.61**	9.76	**41.54**	9.76	**44.43**	9.73	**45.92**	9.12	**47.26**	8.43	**45.30**	9.41
Blood pressure	**0.35**	0.48	**0.35**	0.48	**0.42**	0.49	**0.36**	0.48	**0.56**	0.50	**0.60**	0.49	**0.70**	0.46	**0.59**	0.49
Preventive care	**0.11**	0.31	**0.12**	0.32	**0.13**	0.34	**0.11**	0.32	**0.22**	0.41	**0.20**	0.40	**0.22**	0.42	**0.21**	0.41
Cholesterol	**0.06**	0.25	**0.09**	0.29	**0.12**	0.33	**0.08**	0.27	**0.18**	0.38	**0.27**	0.44	**0.40**	0.49	**0.24**	0.43
Diet and exercise instruction	0.03[#]	0.18	0.03[+]	0.18	**0.06**	0.25	**0.04**	0.19	0.05[#]	0.22	0.05[+]	0.22	**0.13**	0.34	**0.06**	0.24
Specialist	**0.45**	1.79	**0.33**	1.59	0.39[+]	2.08	**0.40**	1.75	**0.96**	2.37	**0.92**	2.71	0.66[+]	1.95	**0.91**	2.46
Other specialist	**1.97**	12.15	**1.76**	9.73	**1.55**	7.72	**1.86**	11.08	**5.84**	19.49	**5.78**	19.95	**4.61**	14.60	**5.68**	19.15
N	3,703		2,084		524		6,311		898		625		202		1,725	

Note: Mean values in bold shows that infrequent users (and within each weight category) differ significantly from frequent users at 0.01 percent level. Mean values with a # shows that infrequent users (and within each weight category) differ significantly from frequent users at 5 percent level. Mean values with a + shows that infrequent users (and within each weight category) differ significantly from frequent users at 10 percent level.

Table A2: Negative binomial model (NBM-1), Two part model (TPM-1). Based on Negative binomial-1. N=8,036

| | NBM-1 | | TPM-1 | | | |
| | | | Part one | | Part two | |
	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
Contracep	0.261***	0.032	0.620***	0.095	0.172***	0.028
Men	-0.501***	0.023	-0.557***	0.040	-0.292***	0.021
Vocational	-0.005	0.025	0.076*	0.043	-0.037*	0.022
Further	-0.050*	0.029	0.023	0.049	-0.068***	0.025
Age	0.001	0.001	-0.005**	0.002	0.004***	0.001
Children	-0.035**	0.016	-0.042	0.027	-0.023*	0.014
Small child	0.150***	0.034	0.208***	0.061	0.092***	0.030
Married	-0.008	0.026	0.101**	0.045	-0.055**	0.023
Immigrant	0.126*	0.073	0.224	0.144	0.049	0.065
Logwage	-0.113***	0.017	-0.134***	0.034	-0.079***	0.015
Emotional	0.106***	0.027	0.192***	0.051	0.042*	0.024
Physical	-0.005	0.021	-0.007	0.036	0.006	0.019
Overweight	0.039*	0.023	-0.015	0.038	0.058***	0.020
Obese	0.115***	0.035	0.048	0.063	0.116***	0.031
Goodhealth	-0.548***	0.026	-0.518***	0.060	-0.426***	0.023
Frkborg	0.077*	0.046	0.048	0.078	0.064	0.040
Roskilde	0.129***	0.044	0.153*	0.079	0.077**	0.039
Vsjaelld	0.048	0.047	0.010	0.081	0.049	0.042
Storst	0.017	0.051	0.024	0.085	0.017	0.045
Bornholm	-0.042	0.061	-0.010	0.102	-0.041	0.054
Fyn	-0.039	0.048	-0.069	0.078	-0.018	0.042
Sjylland	0.067	0.048	0.015	0.081	0.072*	0.043
Ribe	0.081*	0.046	0.108	0.080	0.056	0.040
Vejle	0.023	0.049	0.010	0.081	0.034	0.043
Ringk	0.075	0.047	0.138*	0.082	0.029	0.042
Aarhus	-0.012	0.042	-0.034	0.070	0.005	0.037
Viborg	-0.031	0.047	0.017	0.077	-0.047	0.041
Njylland	-0.031	0.047	-0.036	0.078	-0.030	0.041
Smoke	-0.059**	0.027	-0.132***	0.047	-0.007	0.024
Smore15smo	0.006	0.032	0.045	0.054	-0.017	0.028
Smoked0yr	0.017	0.087	0.128	0.148	-0.035	0.077
Smoked6_10yr	0.022	0.100	0.067	0.170	-0.021	0.088
Smoked11_yr	0.112	0.087	0.205	0.147	0.047	0.076
Smokedmisyr	0.156	0.132	0.254	0.216	0.046	0.116
Constant	3.402***	0.218	3.193***	0.429	2.977***	0.194
Log L	20,451		21,116			

* significant at 10%; ** significant at 5%; *** significant at 1%

Note: Omitted categories are basic education, normal weight, Copenhagen, and smoke1_5yr (smoked 1-5 years)