Smoking habit changes and body weight: causal estimates from the British Household Panel Survey

Pieroni, Luca and Salmasi, Luca

10 December 2012

Smoking habit changes and body weight: causal estimates from the British Household Panel Survey

Pieroni L.*, Salmasi L.**

*University of Perugia (Italy) and University of the West of England (UK)

**University of Perugia (Italy)

Corresponding author: Pieroni L., Department of Economics, Finance and Statistics,

Via Pascoli 20, 06123 Perugia, Italy. E-mail: lpieroni@unipg.it

Abstract

This paper evaluates the causal relationship between smoking and body weight through two waves (2004-2006) of the British Household Panel Survey. We model the effect of changes in smoking habits, such as quitting or reducing, and account for the heterogeneous responses of individuals located at different points of the body mass distribution by quantile regression. We investigate the robustness of our results by means of a large set of control groups and the application of an instrumental variable (IV) estimator. Our results reveal the positive effect of quitting smoking on weight changes, which is also found to increase in the highest quantiles, whereas the decision to reduce smoking does not affect body weight. Lastly, cost-benefit analysis reveals that quitting smoking implies savings for the National Health Service which are much larger than the costs associated with increased obesity.

Keywords: Body Mass Index, Overweight and Obesity, QTE, Instrumental Variable Quantile regression

JEL classification : I10, I12, I18

1 Introduction

Obesity is one of the major risks that individuals face during their life course, particularly in the United States, but it also shows growing importance in Europe. As widely suggested, this epidemic problem may be alleviated by changing some habits at individual level, for example, by decreasing food consumption or promoting physical activity, but other determinants, like smoking, may seriously influence individual results in terms of weight. In fact, the decline in smoking rates, observed in the last few decades, was found

to be associated with higher obesity rates (Chou et al. 2004, Rashad et al. 2006, Flegal 2007, Baum 2009, Liu et al. 2010), although existing empirical studies did not reach a consensus. Gruber & Frakes (2006) reported the opposite of the expected relationship, whereas Nonnemaker et al. (2009) found no significant evidence. This ambiguity arises because changes in smoking habits and weight may also be affected by endogeneity bias deriving from unobserved confounders (Baum 2009), or reverse causality (Cawley et al. 2004), implying that a careful identification strategy must be used to estimate these effects causally.

To estimate the extent to which changes in smoking habits modify individual body weight, expressed in terms of body mass index (BMI), we used a longitudinal dataset extracted from two waves (2004-2006) of the British Household Panel Survey (BHPS), which includes information on smoking and a large number of socio-demographic variables. We exploited the fact that we observed a random sample of the population of smokers in two periods, in which some subjects experienced the transition from smoking to non-smoking, while others reduced their cigarette consumption. The longitudinal framework of the dataset allowed us to separate the effect of quitting or substantially reducing cigarette consumption controlling for unobservable individual time-invariant characteristics.

Our analysis differs from those in the current literature in three important respects. First, we propose a general framework based on a multi-treatment model to test whether a substantial reduction in cigarette consumption can contribute to affecting weight changes. We chose to analyse reductions in smoking because we noted from descriptive analysis that part of the weight variation in obese people was associated with reduced cigarette consumption, rather than only with quitting smoking. Second, we focus on the effects of smoking on different quantiles of BMI distribution. That is, we extend the quantile regression within a difference-in-differences (DID) framework, to measure the quantile treatment effect (QTE). This allows us to evaluate the proposed model across the conditional BMI distribution, taking into account the influence of unobservable characteristics. Third, we evaluate possible heterogeneous effects of quitting or reducing smoking for individuals belonging to different BMI clinical classes, specifying an extension of our model which includes a full set of interactions between treatment and BMI classes. In this way, we estimate the effect of changes in smoking habits on BMI for overweight and obese individuals, which are particularly of interest for policy-makers, and use the estimated

parameters to compare social costs from increased obesity with benefits from quitting smoking.

We propose a strategy which estimates the effects on weight of changes in smoking habits by selecting various ''control groups" in addition to the natural control group of ''smokers", in which each group is able to consider specific confounders. We also propose a sensitivity analysis to investigate to what extent the estimates derived from different identifications are distant from one another.

Our empirical results are the following. First, we document a significant increase in body weight for quitters, particularly if obese. Although point estimates are not very large in magnitude, they increase across BMI distribution. Second, robustness analysis generally confirms these findings when various control groups are used and when an IV estimator is implemented. Third, even though the obese and overweight groups tend to increase less or even decrease BMI in the period considered, they show larger effects from quitting smoking on weight, with respect to normal-weight people. Fourth, results from a social cost-benefit analysis indicate that quitting smoking implies much larger savings in health costs than the costs associated with increased obesity.

In Section 2, we summarise the main features of our empirical framework. We extend the discussion of our strategy from the average treatment effect (ATE) to the quantile treatment effect (QTE), to measure the relationships tested on the conditional BMI distribution. A multi-treatment model of the effect of cigarette reduction on weight changes is proposed, in which quitting is a special case (e.g., single treatment model). In Section 3, we describe our data and focus particularly on DID identification, using a large number of control groups. Section 4 lists our main results and discusses the robustness of the analyses. The policy implications of the cost-benefit calculation are discussed in Section 5, and Section 6 concludes.

2 Modelling the causal effects of smoking on weight changes

2.1 Preliminaries

Our aim was to estimate the causal effect of changes in smoking habits on BMI in a representative sample of the UK population. Let us consider a benchmark model, in which $BMI(i,t)$ is the body mass index of individual i at time t and in which, in a

hypothetical period, some fraction of the population reduces its cigarette consumption (e.g., non-random treatment). That is, individuals are observed in pre-treatment period $t = 0$ and in post-treatment period $t = 1$, in which $D(i,t) = 1$ if an individual has been exposed to the treatment between $t = 0$ and $t = 1$, and $D(i,t) = 0$ if not (control group).

From a theoretical point of view, we assume that ''treated" subjects in $t = 1$ decide to reduce their smoking up to the extreme case of ''zero cigarettes smoked" (i.e., quitting), a situation which is of great interest in the health economics literature, because it is linked with addiction issues (Becker & Murphy 1988, Baltagi & Griffin 2001).

As in Ashenfelter & Card (1985), we assume that the outcome of interest (i.e. BMI) is generated through a component of variance process. A sufficient condition to identify the effect of smoking status changes is that selection for treatment, conditional on covariates, does not depend on individual transitory shocks; that is, $P(D(i,1) = 1|\epsilon(i,t)) = P(D(i,1) = 1)$ for $t = 0, 1$ (Abadie 2005). The model is formally written as:

$$BMI(i,t) = \mu + X(i)'\pi(t) + D(i,1)\eta + \delta t + \beta D(i,t) + \epsilon(i,t) \tag{1}$$

where δ is a time-specific component, β is the effect of treatment, η represents unobserved individual specific characteristics such as various concerns about health, and $X(i)$ is the vector of observed individual characteristics which is assumed to be uncorrelated with individual-transitory shocks $\epsilon(i,t)$, with mean zero at each period and possibly correlated in time.

Provided that D is not endogenous, the empirical specification of equation (1) is given as:

$$E[BMI|D,t,X] = \mu + X(i)'\pi(t) + D(i)\eta + t\delta + (D(i) \times t)\beta \tag{2}$$

in which the OLS estimate of β is the counterpart of the conditional ATE, within the DID approach. In fact, the ATE can be expressed as the difference between the conditional means of treatment and control groups in the two observed periods:

$$\begin{aligned} &E[BMI|D = 1, t = 1, X] - E[BMI|D = 1, t = 0, X] - \\ &E[BMI|D = 0, t = 1, X] - E[BMI|D = 0, t = 0, X] = \beta. \end{aligned} \tag{3}$$

Here we assume that the effect of changes in smoking status may vary according to individual unobserved characteristics, like for example attitude toward risk. The previous model is consequently extended to take into consideration this aspect and a quantile regression estimator is used to obtain the QTE, that is more appropriate than the usual ATE since is able to control for quantile unobserved heterogeneity (Doksum 1974). Formally, the quantile regression model applied to our case is expressed as:

$$q[BMI|D,t,X,\theta] = \mu^{\theta} + X(i)^{'}\pi^{\theta}(t) + D(i,t)\eta^{\theta} + t\delta^{\theta} + (D(i) \times t)\beta^{\theta} \tag{4}$$

where the model parameters are obtained by minimising the weighted sum of residuals[1]. Thus, $q[BMI|D,t,X,\theta]$ is the estimated BMI at the $\theta - th$ conditional quantile and θ is chosen in the interval (0, 1). δ^{θ}, η^{θ}, π^{θ} and β^{θ} are parameters, which now also depend on θ, associated with the already described effects of time, unobservable characteristics, covariates, and the treatment effect. Under the same identifying assumptions used for ATE, the quantile regression estimates of (β^{θ}) in (4) are the DID estimates of the conditional QTE.

2.2 The model

Here, first, we extend the model proposed in the previous section to estimate ATEs and QTEs in a multiple treatment framework. We justify this extension because we want to test whether reducing smoking may also be responsible for weight increases: that is, if smoking reductions are relevant in explaining weight changes, then estimates of the effect of quitting will be biased when individuals who are reducing the number of cigarettes they smoke are also included in control groups.

With respect to quitting, reducing smoking is expected to have a smaller influence in this case, at least because addictive effects are weaker. The empirical evidence shows that many individuals try to quit smoking by gradually reducing the number of cigarettes they smoke each day. This method usually produces minimal or no withdrawal symptoms when the reduction is small, but may cause significant ones when the cut is greater or more sudden. For this reason equation (5) includes the potential effects of reducing smoking, specifying the following multi-treatment model:

[1]See, Koenker & Bassett (1978) and Koenker & Hallock (2001) for a discussion of quantile regression model.

$$BMI(i,t) = \mu + X(i)^{'}\pi(t) + \sum_{j=1}^{n} D_j(i,t)\eta_j + \delta t + \sum_{j=1}^{n}(D_j(i,t) \times t)\beta_j + \epsilon(i,t) \quad (5)$$

where D(j) indicates the $j-th$ treatment group, with $j = 1, ...n$, and the other parameters and variables are those already described. We simplify the model by restricting the possible cases (treatments) to two: i) individuals who quit smoking, i.e., 100% reduction of cigarette consumption, the extreme case; ii) individuals who do not quit smoking but who significantly reduce their cigarette consumption, by at least 50%.

The ATEs for extended model (5) are calculated as the difference between the conditional means of the treatment and control groups in the two observed periods. Formally, the time difference of the conditional means of the treatment group is given as:

$$\begin{aligned} &E[BMI|D = 1, t = 1, X] - E[BMI|D = 1, t = 0, X] \\ &= (\mu + X(i)^{'}\pi + \sum_{j=1}^{n}\beta_j + \sum_{j=1}^{n}\eta_j + \delta) - (\mu + X(i)^{'}\pi + \sum_{j=1}^{n}\eta_j) = \sum_{j=1}^{n}\beta_j + \delta, \end{aligned} \quad (6)$$

while that of the control group is:

$$\begin{aligned} &E[BMI|D = 0, t = 1, X] - E[BMI|D = 0, t = 0, X] \\ &= (\mu + X(i)^{'}\pi + \delta) - (\mu + X(i)^{'}\pi) = \delta. \end{aligned} \quad (7)$$

The result of the double difference is $\sum_{j=1}^{n}\beta_j$, which represents the conditional ATEs to be estimated.

Analogously, we can estimate the multi-treatment model using a quantile regression framework:

$$BMI^{\theta}(i,t) = \mu^{\theta} + X(i)^{'}\pi^{\theta}(t) + \sum_{j=1}^{n} D_j(i,1)\eta_j^{\theta} + \delta^{\theta} t + \sum_{j=1}^{n}(D_j(i,t) \times t)\beta_j^{\theta} + \epsilon^{\theta}(i,t). \quad (8)$$

Based on the same identifying condition, the QTE estimator is consistently identified through DID as:

$$\beta^{\theta*} = \sum_{j=1}^{n} \beta_j^{\theta}. \tag{9}$$

It is worth noting that overweight and obese individuals are of great interest to policy-makers, because preventing weigth excess produces both significant gains in terms of health and reductions in terms of costs for treating their related illnesses. Thus, we consider possible heterogeneous treatment effects, across BMI clinical thresholds, by specifying the following model for the usual four BMI classes allowing for multiple treatment effects:

$$\begin{aligned} BMI_{(i,t,h)} = \mu + X(i)'\pi(t) + \sum_{h=1}^{4}\sum_{j=1}^{2} D_j(i,h)\eta_{j,h} + \sum_{h=1}^{4} t\delta_h \\ + \sum_{h=1}^{4}\sum_{j=1}^{2}(D_j(i,h) \times t)\beta_{j,h} + \epsilon(i,h), \end{aligned} \tag{10}$$

where subscript h ranges from 1, which represents the BMI threshold corresponding to underweight individuals (i.e., BMI < 19), to 4, which refers to the obese category (i.e., BMI ≥ 30)[2]. The OLS estimate of $\beta_{j,h}$ is still the DID estimate of the ATE of treatment j for group h.

The model which only includes quitting smoking as treatment can be derived as a nested specification of (10) under the assumption that smoking reductions do not affect body weight significantly (i.e. $j = 1$):

$$\begin{aligned} BMI_{(i,h)} = \mu + X(i)'\pi + \sum_{h=1}^{4} D(i,h)\eta_h + \sum_{h=1}^{4} t\delta_h \\ + \sum_{h=1}^{4}(D(i,h) \times t)\beta_h + \epsilon(i,h). \end{aligned} \tag{11}$$

[2]Normal-weight and overweight range from $19 - 24.99$ and $25 - 29.99$ of BMI, respectively.

3 Source, identification strategy and descriptive analysis

3.1 Source

The dataset used in this paper was extracted from the BHPS, a multi-purpose survey which reports information at both individual and household level for a representative panel of the UK population. The original sample was drawn from 250 areas of England, and was subsequently enlarged to include Scotland and Wales in 1999 and Northern Ireland in 2002. The dataset is composed of 18 waves and the first survey was conducted in 1980: for our purposes, we use the 14th and 16th waves, conducted respectively in 2004 and 2006, because they also recorded two anthropometric characteristics, height and weight, which allow us to calculate BMI (weight/height^2) for a balanced panel of 13,230 individuals.

3.2 Identification

One difficulty in evaluating the effectiveness of policies aimed at reducing smoking is to verify whether such interventions also have unintended consequences on body weight. Estimating causal parameters is made difficult because the relationship between smoking and BMI is affected by endogeneity due to unobservable characteristics or reverse causality.

To cover these issues, part of the literature uses the introduction of smoking bans as an instrument for smoking, within an IV framework (Liu et al. 2010, Pieroni & Salmasi 2012). Although such anti-smoking policies are expected to affect negatively smoking habits, they are certainly not correlated with the unobserved characteristics affecting body weight.

Another approach proposed in the literature uses a *difference-in-differences* (DID) strategy with panel data. For example, Baum (2009) used the DID estimator to identify the above relationship accounting for individual unobservable time-invariant characteristics affecting cigarette consumption and weight differently for both treatment and control groups. In the present work, in view of the panel nature of our dataset, we also adopt the DID approach and solve endogeneity issues by defining different control groups, aimed at controlling for the bias induced by reverse causality or unobservable characteristics.

In line with the models presented in Section 2, we start our discussion on model identification by defining the following treatment and control groups:

D1. **Treated group of smokers who quit smoking ($\mathbf{TG}_Q$)**: individuals who were smokers in 2004 and became non-smokers in 2006.

D2. **Treated group of smokers who reduced smoking ($\mathbf{TG}_R$)**: individuals who were smokers in 2004 and significantly reduced cigarette consumption (by at least 50%) in 2006.

D3. **Control group of smokers ($\mathbf{CG}_S$)**: individuals who were smokers in 2004 and remained so in 2006.

Our evaluation strategy assumes that weight variations between 2004 and 2006 for TG_Q individuals are affected by quitting smoking and by a spontaneous dynamic (i.e., time-specific component), whereas individuals who continue to smoke (CG_S) are only affected by the spontaneous dynamic. We also evaluate the effect on a treatment group of individuals who significantly[3] reduced their cigarette consumption (TG_R), because these subjects are assumed to be qualitatively exposed to the same symptoms of metabolic rate reduction, withdrawal and changes in eating habits which are responsible for increasing weight, as shown in quitters by Dill et al. (1934), Jacobs et al. (1965), Glauser et al. (1970) and Jacobs & Gottenberg (1981).

However, according to Cawley et al. (2004), smoking habits are influenced by body weight if smokers do not quit because they are afraid of putting on weight. In this case, estimates of the relationship between smoking and body weight are biased by reverse causality. Then, comparing TG_Q or TG_R with CG_S may produce biased estimates because the estimated weight variation for the control group of smokers, CG_S, is biased downwards and consequently the estimated ATEs and QTEs are biased upwards.

We evaluate the magnitude of this effect by defining two control groups. The first also includes non-smokers in the smokers' control group[4], CG_{ALL}, whereas the second includes only non-smokers CG_{NS}. Both groups are composed of individuals who keep their cigarette consumption stable and who, in principle, are not affected by reverse causality, like the CG_S, because their weight does not affect their smoking decisions. We formally define:

[3]We considered as significant a reduction of 50% of current cigarette consumption. We also carried out estimates with thresholds of 30% and 70% for cigarette consumption reductions; the results were very similar to those obtained with the 50% threshold. All tables related to these estimates are available from the authors upon request.

[4]We excluded from this control group non-smokers who started smoking in 2006, irrespective of whether they were or were not ex-smokers in 2004. We anticipate that the dimension within our sample (2.04%) is negligible for our estimates.

D4. **Control group of smokers and non-smokers ($\mathbf{CG}_{ALL}$)**: individuals who were smokers or non-smokers in 2004 and remained so in 2006.

D5. **Control group of non-smokers($\mathbf{CG}_{NS}$)**: individuals who were non-smokers in 2004 and remained so in 2006.

Since treatment is not randomly assigned in our dataset, treated and control groups may also differ according to time-varying unobserved factors related to smoking and weight decisions. In this case, the presence of unobserved heterogeneity may cause our baseline estimates, comparing treatment groups (i.e., TG_Q or TG_R) with CG_S, to be biased downwards. In fact, quitters are generally more concerned about their health and more oriented towards the future, as discussed by McCaul et al. (2006) and their decision to quit smoking may thus be seen as part of a more general attitude aimed at improving health - for example, by also reducing weight in obese people. The presence of these individuals in TG_Q or TG_R may bias the estimated weight variation as well as the estimated ATEs downwards.

In order to take into consideration this aspect, we made up a new control group composed of individuals who were smokers in 2004 and 2006, but who quit in 2008. That is, the BMI variation of ''next period" quitters was considered as the most appropriate control group for the BMI variation of TG_Q, because they have the most similar unobservable characteristics according to future health behaviour. A similar strategy was proposed and applied to the job market by Del Bono & Vuri (2011). Formally, our new control group is defined as:

D6. **Control group of ''next period" quitters (CG_{Q08})**: individuals who were smokers in 2004 and 2006, but who quit in 2008.

Analogously we used the same argument as above to define a control group of individuals who reduced smoking in 2008, defined formally as:

D7. **Control group of ''next period" reducers (CG_{R08})**: individuals who were smokers in 2004 and 2006, but who substantially reduced smoking in 2008.

In the next sub-section, we empirically justify the use of both the QTE estimator, and the ATE to estimate the effects on weight changes of quitting or reducing cigarette

consumption.

3.3 Preliminary analyses and covariates

To emphasize the differences between treated and control groups at different quantiles of the outcome variable, we compared the estimated densities of BMI empirical distributions, shown in Figure 1, of ''non-smokers" (solid line) and ''heavy smokers"[5] (dashed line). The estimated empirical distributions are both skewed and non-normally distributed and also have different shapes. The empirical distribution of non-smokers is shifted to the right with respect to that of heavy smokers, implying that the former group also has higher BMI. This result is consistent with the hypothesis that changes in cigarette consumption (reducing or quitting smoking) may increase overweight and obesity rates.

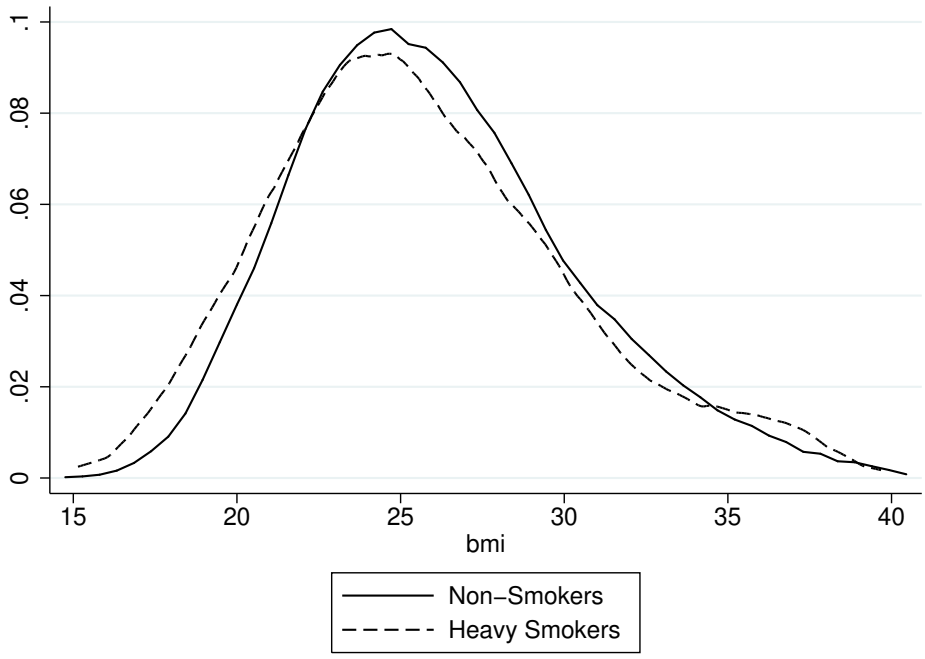

Figure 1: BMI kernel density estimate, by smoking status

Table 1 (panel a) shows an average BMI increase of 0.85 points for TG_Q and a smaller variation for CG_S (0.24 BMI points); the unconditional ATE, according to these two groups, is therefore estimated to be 0.61 BMI points. Also according to other control groups, the estimated ATE is about 0.6 BMI points, except for CG_{Q08}, where it is slightly larger. Panel (b) lists the effects of reducing smoking (TG_R) on BMI; we estimate an

[5]We consider adults who currently smoke more than 20 cigarettes per day as "heavy smokers".

ATE very close to zero (0.01) for this group. The same result holds for all control groups.

We had previously checked that BMI differences between smokers and non-smokers were different over the empirical distribution. We then used the usual BMI clinical thresholds to calculate unconditional ATEs for underweight, normal-weight, overweight and obese individuals. Table 2 shows two important results. First, the effect of quitting smoking on BMI is greater for obese individuals (range: 0.97 to 1.08 points), irrespective of control group used. Second, when we examine the effect of reducing cigarette consumption, we find a negative, although small, effect on obese people (range: $-.21$ to $-.26$ points), compensated by a (small) positive effect on normal-weight individuals. At this stage, however, we cannot comment on the significance of these effects and leave the discussion to the next sections.

Table 1: BMI absolute variations and ATEs (2004-2006), by smoking status group

	N. obs	**BMI**	**BMI**	**Absolute variation**	**ATE**
Panel a		**2004**	**2006**	**2004/2006**	**2004/2006**
TG_Q	443	25.04	25.89	0.85	-
CG_S	2455	25.12	25.36	0.24	0.61
CG_{ALL}	10946	25.82	26.06	0.25	0.6
CG_{NS}	8491	26.02	26.27	0.25	0.6
CG_{Q08}	335	24.6	24.78	0.18	0.67
Panel b					
TG_R	694	25.02	25.3	0.28	-
CG_S	1761	25.14	25.42	0.28	0
CG_{ALL}	10252	25.82	26.09	0.27	0.01
CG_{NS}	8491	26	26.25	0.25	0.03
CG_{Q08}	1011	25.46	25.69	0.23	0.05

Notes: TG_Q and TG_R: treated groups of quitters and individuals reducing consumption of cigarettes by more than 50%, respectively. CG_S, CG_{ALL}, CG_{NS} and CG_{08}: control groups (see subsection 3.2.)

We also include a full set of covariates, both time invariant and time varying, in our model to control for observable individual characteristics. All time-varying variables are introduced into the models as variations between waves, as in French et al. (2010). The matrix X of covariates is composed of: gender; health status (five modalities, from excellent to very poor); health status variations (positive variations, from very poor or

Table 2: BMI absolute variations and ATEs (2004-2006), by smoking status group and BMI class

	Absolute variation				**ATE**			
BMI classes	< 19	$19 - 25$	$25 - 30$	> 30	< 19	$19 - 25$	$25 - 30$	> 30
TG_Q	1.15	0.71	1	0.78	-	-	-	
CG_S	0.35	0.41	0.2	-0.3	0.81	0.3	0.8	1.08
CG_{ALL}	0.4	0.42	0.23	-0.21	0.75	0.29	0.76	0.99
CG_{NS}	0.44	0.42	0.24	-0.19	0.72	0.29	0.76	0.97
CG_{Q08}	0.41	0.36	-0.01	-0.19	0.74	0.35	1.01	0.97
TG_R	0.51	0.41	0.28	-0.4	-	-	-	
CG_S	0.34	0.46	0.15	-0.43	0.16	0.25	0.13	0.03
CG_{ALL}	0.43	0.44	0.25	-0.18	0.08	0.28	0.03	-0.21
CG_{NS}	0.46	0.43	0.27	-0.13	0.05	0.28	0.01	-0.26
CG_{Q08}	0.29	0.4	0.19	-0.15	0.22	0.31	0.09	-0.25

Notes: TG_Q and TG_R: treated groups of quitters and individuals reducing consumption of cigarettes by more than 50%. CG_S, CG_{ALL}, CG_{NS} and CG_{08}: control groups (see, subsection 3.2). BMI clinical classes: < 19, underweight; $19 - 25$, normal-weight; $25 - 30$, overweight; > 30, obese.

poor to good or excellent; and negative variations, the opposite); physical activity (at least once a week or less); physical activity variations (positive and negative, from at least once a week to less or the opposite); length of sickness (long-term versus short-term or healthy); length of sickness variations (from long-term to short-term or healthy, or the opposite); work conditions (weekly hours of work; strenuousness of job; presence in the household of a working mother); variations in working conditions (changes in number of weekly hours of work or job strenuousness); ethnic group membership (white or other ethnicity); age (five classes, 18-29, 30-39, 40-49, 50-59, ≥ 60); net income (five classes, by quintile of net income distribution); marital status (married; couple; divorced; separated; widowed); marital status variations (from married, couple, divorced, separated, widowed, to another category); education (degree; diploma; GCE A-level and/or O-level examinations)[6]; and three country dummies (England, Wales, Scotland). We also include a measure of alcohol consumption (five modalities of frequency with which the respondent drinks outside the home, from ''at least once a week" to ''never") which is considered as a proxy for unobserved individual heterogeneity in addictive behaviour, since alcohol is known to be a complement to smoking, at least in the Italian case (Aristei & Pieroni

[6]A-levels and O-levels refer to the examinations for the General Certificate of Education offered by educational institutions in the United Kingdom and a few of the former British colonies. In particular, A-levels were the subsequent examinations for those who studied for a further two years after O-levels at the age of 16.

2010). Time varying covariates are used in our specification to control for the effect of various life-style shocks, mostly health-related, which could also have an effect on the decision to quit smoking and may influence weight as well. In addition we included a set of dummy variables recording whether subjects were underweight, normal-weight, overweight or obese in 2004, and a full set of interactions with the time trend to account for the different BMI dynamics observed in Table 2. These covariates are included as proxies of health concerns or orientation towards the future of individuals with different initial conditions in terms of BMI. Summary statistics of covariates for the treated (TG_Q) and smoker control groups (CG_S) are listed in Appendix A.

4 Results

4.1 OLS estimates

In this section, we discuss the OLS estimates of the effect of quitting (e.g., single treatment using TG_Q) and both quitting or reducing (e.g., multi-treatment using TG_Q and TG_R) on weight changes, with smokers (CG_S) as the control group. Table 3 lists the estimated coefficients. The parameters of the model with single treatment (first column of the upper part of Table 3 are in line with the unconditional ATEs shown in Table 2; quitting smoking leads to a small but significant increase in terms of BMI of about 0.59 (s.e.=0.100). Moreover, there are not significant differences in terms of BMI between treated and control groups, since η_1 is not statistically different from zero. This result is not surprising because, under this specification, we are comparing groups of individuals, who in 2004 were smokers and had similar initial conditions in terms of BMI. Instead, spontaneous dynamic δ is positive and significant.

When we compare these results with those obtained with the other control groups, CG_{NS} and CG_{ALL} (Table 3), we find that the effect of quitting is slightly smaller than that obtained from previous estimates, and is respectively 0.53 (s.e. = 0.095) and 0.54 (s.e. = 0.094). As expected the effect of quitting smoking on BMI estimated from these control groups, which account for the effect of reverse causality, is lower than that estimated through CG_S. However, we can show that these differences are negligible in terms of weight if we consider that the weight increase for an individual at the sample average - with height of 1.70 m - who decides to quit smoking is 1.7 kg under CG_S, whereas if we

consider CG_{NS} or CG_{ALL} the estimated weight increases are respectively 1.5 kg and 1.6 kg.

Also with CG_{NS} or CG_{ALL} control groups BMI has a positive and significant trend. The only difference with CG_S worth mentioning regard the η_1 coefficients, which now are negative and significant. This result is in line with findings from the medical literature, according to Grunberg (1985), Klesges et al. (1989) and French & Jeffery (1995), who found that smokers weighed less than non-smokers because smoking increases the metabolic rate and thus the number of calories consumed by the body during the day: after one cigarette is smoked, the heart may beat 10-20 times more than its normal rate per minute, (Dill et al. 1934, Hiestand et al. 1940, Glauser et al. 1970), and the amount of weight gained after quitting is usually close to the initial gap between smokers and non-smokers, at least in the short term (Keys et al. 1966, Karvonen et al. 1959, Higgins 1967).

Table 3 also shows the estimated coefficients obtained with CG_{08} as the control group. As expected, "next period" quitters have equivalent initial conditions to those of current smokers, as the non-significant η parameter shows. However, the ATE estimates (β_1), with this control group, are higher than the previous specifications and reach 0.63 (s.e. = 0.142). In line with our expectation, this control group seems to be able to correct the downward bias induced by unobservable variables, although the difference with respect to our baseline estimates is quite limited, especially when considered in terms of weight. In this case the estimated weight variation for a representative individual at the sample mean is of 1.8 kg and is, as well as for the previous control groups, close to the weight variation estimated with CG_S.

The lower part of Table 3 lists parameters associated with the effects of both quitting and reducing smoking (i.e., equation (5)). The estimated β_1 and δ are close to those of the upper part of the table, whereas the coefficient linked with smoking reduction, β_2, is non-significant.

4.2 Quantile estimates

In this section we examine the possibility that the effect of quitting smoking varies across the conditional BMI distribution using the quantile regression model. Figure 2 shows the scatter plot of unconditional versus conditional BMI distribution values, obtained from

Table 3: Causal effect of cigarette consumption changes on BMI (2004-2006) by control group

a	**Treatment: quitting smoking**			
Variables	CG_S	CG_{NS}	CG_{ALL}	CG_{08}
η_1	-0.21	-0.29***	-0.25***	0.03
	(0.181)	(0.077)	(0.075)	(0.105)
β_1	0.59***	0.53***	0.54***	0.63***
	(0.100)	(0.095)	(0.094)	(0.142)
δ	0.31***	0.35***	0.36***	0.35**
	(0.064)	(0.033)	(0.030)	(0.158)
Constant	22.26***	22.70***	22.52***	22.16***
	(0.228)	(0.177)	(0.142)	(0.590)
Observations	5,796	17,868	22,778	1,656
R-squared	0.80	0.80	0.80	0.78
Adj. R-squared	0.80	0.80	0.80	0.77
b	**Multi-treatment: quitting or reducing smoking**			
η_1	-0.17	-0.29***	-0.25***	-0.15
	(0.184)	(0.077)	(0.075)	(0.192)
η_2	0.10	-0.08	-0.02	0.12
	(0.069)	(0.062)	(0.061)	(0.079)
β_1	0.61***	0.53***	0.54***	0.63***
	(0.103)	(0.095)	(0.094)	(0.111)
β_2	0.05	-0.03	-0.01	0.07
	(0.078)	(0.069)	(0.068)	(0.088)
δ	0.30***	0.36***	0.36***	0.30***
	(0.067)	(0.032)	(0.030)	(0.088)
Constant	22.22***	22.70***	22.52***	22.23***
	(0.228)	(0.162)	(0.143)	(0.283)
Observations	5,796	19,270	22,778	4,206
R-squared	0.80	0.80	0.80	0.80
Adj. R-squared	0.80	0.80	0.80	0.80

*Notes: panel (a) lists estimates of BMI model in equation (5), restricted to effects of quitting smoking (e.g., $j = 1$); panel (b) in a multi-treatment framework, shows effects of reduction of number of cigarettes by more than 50%, in addition to quitters (e.g., $j = 2$). CG_{ALL}, CG_{NS} and CG_{08}: control groups described in subsection 3.2. Standard errors in brackets; significant levels as follows notation: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

the fitted values of our baseline specification which uses CG_S. Much of the BMI variability remains unexplained, even after controlling for observable individual characteristics, and it therefore seems reasonable to assume that individuals with different BMI levels,

conditional on observable characteristics, respond differently to changes in smoking.

Figure 2 also shows how the quantiles of unconditional and conditional BMI distributions are different: for example, although a large share of overweight and obese individuals are located above the upper percentiles of both distributions, we cannot directly interpret quantile parameters (i.e., β_1^θ) as the causal effect of quitting smoking on overweight and obesity but rather as the treatment effect across unobservable individual characteristics. As done by Kowalski (2009), in order to give an interpretation of quantile regression coefficients we must make assumptions about the nature and dimension of the unobservable component in our model. Here, we assume that the main source of unobserved heterogeneity is given by individuals' attitude toward risk. We assume that more risk-prone individuals will gain more weight, because they will not care about the possible adverse consequences of quitting on their weight.

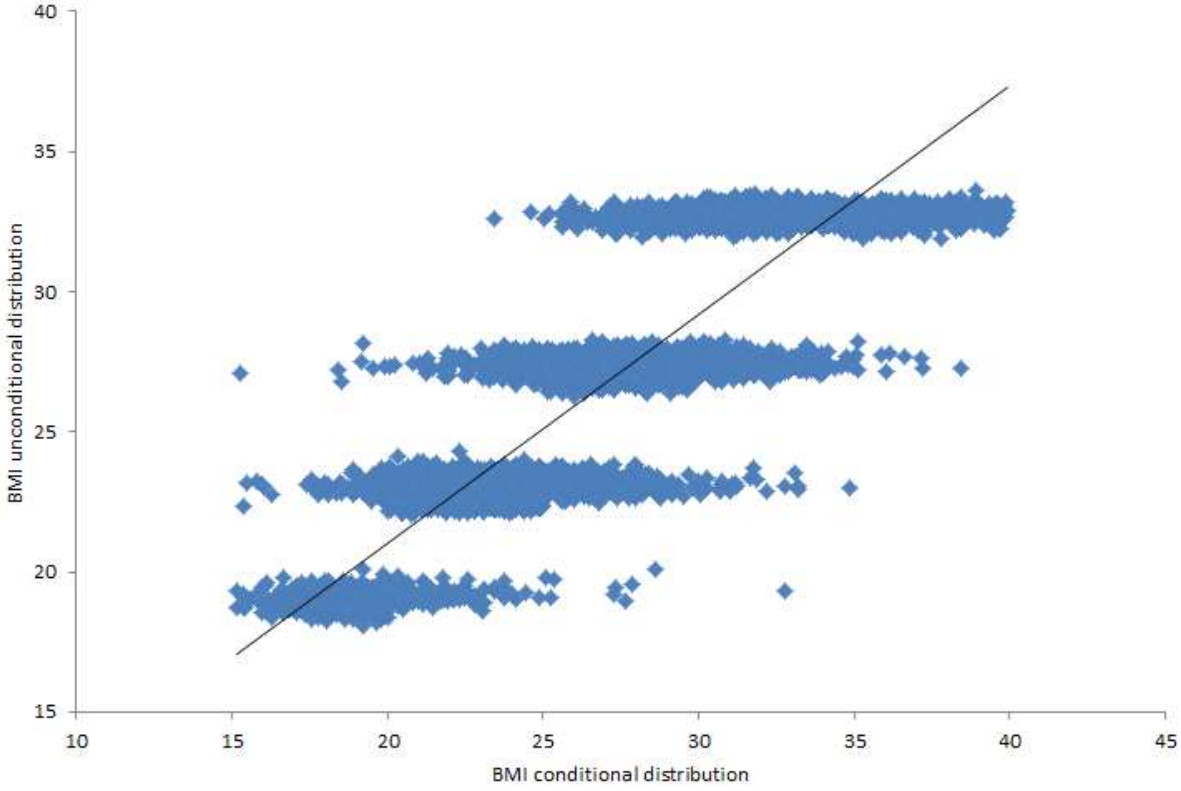

Figure 2: Scatter plot of BMI unconditional versus BMI conditional distribution values

Panel a) of Table 4 lists the QTEs of quitting smoking on BMI. The estimated QTEs highlight a significant and increasing effect across the conditional BMI distribution and vary from a non-significant effect at the 10-th percentile to 0.44 (s.e.=0.178) at the 25-

th, 0.74 (s.e. = 0.199) at the 50-th, 0.70 (s.e.=0.233) at the 75-th and 0.81 (s.e. = 0.261) at the 90-th percentile of the conditional BMI distribution. If we interpret these coefficients as the effect of quitting smoking across the distribution of the unobserved BMI heterogeneity represented by risk propensity, we can conclude that individuals with lower propensity toward risk tend to increase their BMI less after quitting, because they are more careful about their weight, whereas more risk-prone individuals will probably put on more weight - for example, by substituting smoking with more calorie-rich food or not increasing physical activity. Panel b) of Table 4 lists the estimates from the multi-treatment model. Irrespective of the quantile analysed we find that, even in this case, BMI variations are entirely determined by quitting smoking, with very similar coefficients with respect to the single treatment model.

4.3 Heterogeneous effects across BMI clinical thresholds

One of the objectives of our paper is to estimate the weight gained by obese people after quitting smoking. We do this, by interacting individuals belonging to a BMI class in 2004 of underweight, normal-weight, overweight and obese with treatments. That is, we can test whether quitting or reducing smoking have heterogeneous effects on BMI variations in individuals with different initial conditions in terms of BMI. For the sake of simplicity we show here only estimates of the effect of quitting smoking, since even in this specification the effect of reductions is never significant[7].

Table 5 lists OLS estimates, in which smokers are assumed to be the control group (CG_S). For the reference category (normal-weight individuals, $h = 2$), we estimate an effect of quitting smoking on BMI of 0.31 (s.e. = 0.122) and this coefficient is found to increase for individuals who belong to higher BMI clinical categories. For overweight individuals ($h = 3$), quitting smoking produces an extra BMI increase, with respect to the effect for the reference category, of 0.48 points (s.e. = 0.214), so that the total BMI increase for these individuals is 0.79. When we interact the treatment with the obese category, we also find a positive and significant coefficient of 0.75 (s.e. = 0.384), implying a total BMI variation of 1.06 points. Lastly, underweight people present a non-significant coefficient, meaning that their BMI variation after quitting is not statistically different from that of normal-weight people.

[7]These estimates are available from the authors upon request.

Table 4: Causal effect of quitting smoking on BMI, quantile estimates with CG_S

a	**Treatment: quitting smoking**				
Variables	**Q10**	**Q25**	**Q50**	**Q75**	**Q90**
η_1	0.05	0.04	0.04	0.04	0.03
	(0.080)	(0.098)	(0.127)	(0.110)	(0.105)
β_1	0.11	0.44**	0.74***	0.70***	0.81***
	(0.161)	(0.178)	(0.199)	(0.233)	(0.261)
δ	-0.06	0.14	0.17	0.33***	0.89***
	(0.127)	(0.107)	(0.122)	(0.111)	(0.157)
Constant	20.03***	21.01***	22.41***	23.77***	24.27***
	(0.335)	(0.224)	(0.286)	(0.251)	(0.245)
Observations	5,796	5,796	5,796	5,796	5,796
R-squared					
Adj. R-squared	.	.	.	.	.
b	**Multi-treatment: quitting or reducing smoking**				
η_1	0.05	0.05	0.04	0.04	0.03
	(0.099)	(0.099)	(0.121)	(0.094)	(0.103)
η_2	0.08	0.07	0.06	0.09	0.06
	(0.093)	(0.099)	(0.081)	(0.079)	(0.062)
β_1	0.16	0.44***	0.77***	0.74**	0.77**
	(0.221)	(0.154)	(0.164)	(0.343)	(0.323)
β_2	0.11	0.06	0.09	0.11	0.01
	(0.180)	(0.154)	(0.070)	(0.143)	(0.204)
δ	-0.07	0.11	0.14	0.30**	0.91***
	(0.171)	(0.111)	(0.131)	(0.137)	(0.183)
Constant	20.03***	21.02***	22.35***	23.73***	24.25***
	(0.226)	(0.171)	(0.258)	(0.157)	(0.176)
Observations	5,796	5,796	5,796	5,796	5,796
R-squared					
Adj. R-squared	.	.	.	.	.

*Notes: panel (a) lists estimates of BMI model in equation (8), restricted to effects of quitting smoking (e.g., $j = 1$); panel (b) in a multi-treatment framework, shows effects of reduction of number of cigarettes by more than 50%, in addition to quitters (e.g., $j = 2$). Standard errors in brackets; significant levels as follows notation: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

If we look at the δ coefficients, we find a positive and significant variation for normal-weight people of 0.36 (s.e. = 0.065). Overweight and obese individuals present lower trends with respect to the reference category, respectively of -0.20 (s.e. = 0.079) and -0.71 (s.e. = 0.14). In order to calculate the trends for the two categories, we must

add to the reference category's parameter the effects estimated through the interaction terms. That is, the BMI of overweight and obese individuals varies by 0.16 and -0.35 points respectively between 2004 and 2006, consistently with the results from descriptive statistics shown in Table 2.

Table 5 also lists the estimated parameters from other control groups. We find that the ATEs obtained when we use CG_{NS} and CG_{ALL} as alternative control groups are lower, but not very distant from those obtained from our baseline model. Lastly, we also show the estimates using CG_{08} as control group. Although the ATE is in line with our expectations in terms of magnitude, we find that is not statistically different from zero. However, the small size of the obese subsample when CG_{08} is used as control group does not allow us to make accurate inferences[8].

4.4 IV estimates

We now present a sensitivity analysis of our estimates to check whether the proposed approach is successful in producing robust results. In the previous sections, we defined a set of control groups in order to verify our baseline estimates with respect to the possible bias induced by endogeneity. Alternatively we can use an IV strategy, which consists of finding one or more exogenous instruments ($Z(i)$) which must satisfy two properties: they must be strong predictors of the endogenous regressor - in our case, quitting smoking - and they must not be correlated with the unobservable component of BMI. In other words, we need to identify an exogenous variation in smoking behaviour induced by the instruments, which is not correlated with the unobservable characteristics of BMI and smoking.

We use as instrument the percentage of smokers in the same socio-economic group of respondents which is a measure of the effect of social interactions on smoking habits. The underlying idea is that the utility that one smoker receives from consuming a cigarette is increased by that of other individuals with the same habit and similar socio-economic characteristics. Specifically, we calculate the percentage of smokers in each region according to age, income, education, occupation and marital status. When constructing this indicator, we exclude from calculation the respondents cigarette smoking. This instrument is assumed to affect the decision to quit smoking negatively and to be uncorrelated with changes in BMI.

[8]With CG_{08} as control group, there are 180 obese individuals in treatment and control groups.

Table 5: Causal effect of quitting smoking on BMI, by clinical classes and control group

a	**Treatment model: quitting smoking**			
Variables	CG_S	CG_{NS}	CG_{ALL}	CG_{08}
η_1	-0.09	-0.27***	-0.21**	-0.04
	(0.107)	(0.101)	(0.099)	(0.149)
$\eta_{1,1}$	0.03	-0.10	-0.02	-0.21
	(0.241)	(0.233)	(0.230)	(0.299)
$\eta_{1,3}$	-0.20	-0.04	-0.09	0.15
	(0.163)	(0.154)	(0.152)	(0.220)
$\eta_{1,4}$	-0.43	-0.04	-0.13	0.36
	(0.387)	(0.363)	(0.361)	(0.500)
β_1	0.31**	0.26**	0.27**	0.32*
	(0.122)	(0.115)	(0.115)	(0.182)
$\beta_{1,1}$	0.47	0.41	0.45	0.32
	(0.470)	(0.466)	(0.463)	(0.514)
$\beta_{1,3}$	0.48**	0.47**	0.47**	0.65**
	(0.214)	(0.202)	(0.201)	(0.310)
$\beta_{1,4}$	0.75*	0.63*	0.65*	0.53
	(0.384)	(0.360)	(0.359)	(0.590)
δ	0.36***	0.37***	0.37***	0.36**
	(0.065)	(0.034)	(0.030)	(0.173)
δ_1	-0.07	0.02	-0.01	0.04
	(0.098)	(0.074)	(0.058)	(0.232)
δ_3	-0.20***	-0.17***	-0.18***	-0.37
	(0.079)	(0.038)	(0.034)	(0.233)
δ_4	-0.71***	-0.60***	-0.62***	-0.49
	(0.140)	(0.061)	(0.056)	(0.459)
Constant	22.25***	22.70***	22.52***	22.23***
	(0.230)	(0.178)	(0.143)	(0.608)
Observations	5,796	17,868	22,778	1,656
R-squared	0.80	0.80	0.80	0.78
Adj. R-squared	0.80	0.80	0.80	0.78

*Notes: panel (a) lists estimates of BMI model in equation (11), restricted to effects of quitting smoking (e.g., $j = 1$); panel (b) in a multi-treatment framework, shows, effects of reduction of number of cigarettes by more than 50%, in addition to quitters (e.g., $j = 2$). $h = 1, 2, 3, 4$ represents BMI clinical classes of underweight, normal-weight, overweight and obesity. Normal-weight is the reference modality. Standard errors in brackets; significant levels as follows notation: p-value *** $\leq$ 0.01, ** $\leq$ 0.05, * $\leq$ 0.1.*

In order to perform IV regression, we estimated the single treatment model presented in equation (5) as its analogous individual first-difference model. Table 6 lists the estimated

coefficients. We find a significant effect for quitters of 0.71 (s.e. = 0.148) and note that this value is not much higher than that obtained by OLS in the baseline model (i.e., 0.59). In the obese subsample, we find that the IV estimate is 1.34 (s.e = 0.614) and is again slightly larger in magnitude than in the OLS estimates (i.e., 1.06, see Table 5). Although these results show that endogeneity biases our baseline estimates (CG_S) downwards, together with those which use CG_{NS} and CG_{ALL}, the difference with respect to the estimates given by the IV approach are not large in terms of weight. If we consider a representative individual at the sample average, the estimated weight gain with IV parameters is 2.05 kg, whereas for a representative obese individual is 3.87 kg, which are not distant from the corresponding OLS estimates of 1.5 kg and 3.1 kg, respectively.

Table 6: Causal effect of quitting smoking on BMI: IV estimates, with CG_S

Variables	**Full sample**	**Obese**
η_1	-	-
β_1	0.71***	1.34**
	(0.148)	(0.614)
δ	0.28***	-0.40*
	(0.060)	(0.224)
Observations	2,898	379
R-squared	0.03	0.04
Adj. R-squared	0.03	0.01

*Notes: Standard errors in brackets. Significant levels as follows: p-value *** ≤ 0.01, ** ≤ 0.05, * ≤ 0.1.*

5 Social costs and benefits

Our estimates identify the robust negative causal effect of smoking on body weight, which is larger in magnitude for overweight and obese people. This result is of crucial importance in the health economic literature, since policy-makers are interested in understanding whether the benefits from savings deriving from reduced smoking are larger or smaller than the extra costs generated by increased obesity.

Panel a) of Table 7 shows the NHS costs associated with smoking in 2006, which were estimated by Allender et al. (2009) to be £5,170 million for the UK. As 22% of the

UK population (60,587,600 in 2006) were smokers, we can calculate the per capita cost of smoking of £388. In addition, using the estimated total direct cost attributable to obesity of £3,532 million (McCormick & Stone 2007) and considering that the obese represented the 24% of the UK total population in 2006, we obtain a per-capita cost of obesity of £243.

Panel b) of Table 7 lists the benefits of reduced smoking and the costs of the resulting increased obesity. From the BHPS, we can estimate that obese smokers decreased by 2.78% between 2004 and 2006, so that this percentage can be used to calculate that the direct social benefits, in terms of NHS cost reductions attributable to smoking, were £156.81 million. This figure is obtained by multiplying the number of obese individuals who quit smoking (i.e., 406,000, 2.8% of 14.54 million obese individuals in 2004) by the per capita cost of smoking (£388).

We also calculated the estimated social costs associated with increased obesity, attributable to quitting smoking. They arise from two different sources: 1) the extra costs of also having to treat those individuals who were overweight in 2004 and who, because of smoking-related weight increases, became obese in 2006; 2) the ''lost benefits" for having to treat those individuals who were already obese in 2004 and who, because they stopped smoking, did not become overweight in 2006. From our empirical model, the BMI in 2006 is a linear function of two components, $\delta+\beta_1$, all other characteristics remaining constant. Estimates of equation (11) listed in Table 5 show the effect of quitting smoking on BMI of overweight individuals in 2004, which is $\beta_1 + \beta_{1,3} = 0.79$. In order to exclude those individuals who would have become obese because of other factors, we must add to this value the estimated BMI growth trend, represented by $\delta + \delta_3$, of 0.16. In other words, our model predicts that overweight smokers who quit smoking and had a BMI greater than 29.05 (but lower than 29.85) in 2004, will become obese in 2006 (BMI > 30 in 2006) because of quitting smoking. This quota of individuals represents an extra cost which the NHS would not have sustained in a scenario in which nobody quits.

In addition, again looking at Table 5, our estimates indicate a negative trend for obese people ($\delta + \delta_4 = -0.35$), implying that those who had a BMI higher than 30, but lower than 30.35 in 2004, would become overweight in 2006 and do not represent a cost (at least in terms of obesity) for the NHS. But, if we look at the obese individuals who also quit smoking in 2006, the positive effect associated with quitting, $\beta_1 + \beta_{1,4} = 1.06$, clearly

overcompensates the natural decreasing trend of BMI, and keeps a significant number of individuals trapped in obesity. These people represent a lost benefit for the NHS in terms of savings due to reduced obesity. All the other individuals classified as obese in 2006 would have been obese in any case, and consequently are not considered as a burden for the NHS.

Again panel b) of Table 7, shows the estimates of the costs described previously. The percentage of individuals with BMI > 29.05 and < 29.85 in 2004, estimated by the BHPS, is 4.43% of the total population (i.e., 2.68 million people). If we assume that 2.78% of the UK population (i.e., a sample of 74,619) quit smoking in 2006, we obtain an additional cost for the NHS of £18.13 million (74,619 × £243). Moreover, the percentage of individuals with BMI > 30 and < 30.35 in 2004 was 1.71% (i.e., 1,036 million people). Also in this case, we assume that the 2.78% of these individuals quit smoking (i.e., 28,803 people) and estimate a lost benefit of £6.99 million (20.803 × £243). Lastly, the net benefit of £131.69 million is estimated by subtracting these costs from the total benefit.

Table 7: Social costs and benefits of quitting smoking

Panel a		
UK Population	60.59	Millions
Percentage of smokers	0.22	%
Cost of smoking	5170.5	Millions of £
Percentage of overweight people	0.38	%
Percentage of obese people	0.24	%
Cost of obesity	3532	Millions of £
Per capita cost of smoking	387.91	£
Per capita cost of obesity	242.9	£
Social benefits of quitting smoking	156.81	Millions of £
Panel b		
$(\delta + \delta_3) + (\beta_1 + \beta_{1,3})$, CG_S	0.79+0.16	
$(\delta + \delta_4) + (\beta_1 + \beta_{1,4})$, CG_S	1.06-0.35	
Obese quitters between 2004/2006 (estimated by BHPS)	0.0278	%
Percentage of overweight individuals with BMI > 29.05 and < 29.85	4.43	%
Percentage of obese individuals with BMI < 30.35 and > 30	1.71	%
Social costs of increased obesity: overweight with BMI >29.05 and < 29.85	18.13	Millions of £
Social costs of increased obesity: obese with BMI < 30.35	6.99	Millions of £
Net social benefits of quitting smoking	131.69	Millions of £

6 Conclusions

Much has been written about the determinants of weight gains in modern society, with particular emphasis on the health risks of obese people. Among these, abandoning the habit of smoking appears to play a significant role in view of the close correlation between the number of those who stop smoking and their body weight. This paper extends the current literature to examine whether a trade-off exists between smoking and weight gains using longitudinal data from the UK. Identification of causal coefficients in terms of BMI changes was carried out in a framework which included the potential effect of people who reduced smoking as well as quitting. In order to consider the heterogeneous response of BMI to changes in cigarette consumption, we also used quantile regression to obtain QTE.

We reject the hypothesis that reduced smoking can affect weight changes at each estimated point of the BMI distribution, as an increasing effect was revealed in obese people who quit. Robustness analysis confirmed these findings when several control groups were used. We also show the slightly greater impact of quitters on weight changes when the model parameters were estimated with the IV estimator. The latter result becomes evident around the threshold of obesity, indicating that most smokers who quit were concerned about their health, although this may imply a negative externality in terms of weight.

Following on from these results, we also present cost-benefit calculations, in which the benefits of quitting in the UK are greater than the cost of overweight. These estimates should be carefully and cautiously interpreted. First, we observe the heterogeneous behaviour of quitting around the obese cut-off point. In our cost-benefit analysis, lack of information of various points of the BMI distribution only allows us to use, for example, the mean of the obese group, with a plausible bias as regard the exact cost of giving up smoking. Secondly, we assess the effect of quitting on BMI over a period of two years, which is perhaps not long enough for complete evaluation of the effects, as pointed out in the meta-analyses of Klesges et al. (1989) or by Courtemanche (2009). However, even after these criticisms have been taken into account, the estimate of the net social benefit, which emerges from our calculations seems large, indicating that continuing to emphasise the health risks of smoking should be maintained and that current smokers should be encouraged to abandon the habit, diminishing the social cost of correlated diseases.

Acknowledgements

This paper is based on work carried out during a visit to the European Centre for Analysis in the Social Sciences (ECASS) at the Institute for Social and Economic Research, University of Essex, supported by the ''Access to Research Infrastructures action under the EU Improving Human Potential Programme".

Data from the British Household Panel Survey (BHPS) were supplied by the UK Data Archive. Neither the original collectors of the data nor archive personnel have any responsibility for the analyses or interpretations presented here.

We thank Amanda Kowalski and the participants to the ''4th Annual Health Econometrics Workshop" at Hunter College of The City University of New York for their valuable comments and suggestions.

References

Abadie, A. (2005), 'Semiparametric difference-in-differences estimators', *Review of Economic Studies* **72**(1), 1–19.

Allender, S., Balakrishnan, R., Scarborough, P., Webster, P. & Rayner, M. (2009), 'The burden of smoking-related ill health in the UK', *Tobacco Control* **18**(4), 262–7.

Aristei, D. & Pieroni, L. (2010), 'Habits, complementarities and heterogeneity in alcohol and tobacco demand: A multivariate dynamic model', *Oxford Bulletin of Economics and Statistics* **72**(4), 428–457.

Ashenfelter, O. & Card, D. (1985), 'Using the longitudinal structure of earnings to estimate the effect of training programs', *The Review of Economics and Statistics* **67**(4), 648–60.

Baltagi, B. H. & Griffin, J. M. (2001), 'The econometrics of rational addiction: The case of cigarettes', *Journal of Business & Economic Statistics* **19**(4), 449–54.

Baum, C. L. (2009), 'The effects of cigarette costs on BMI and obesity', *Health Economics* **18**(1), 3–19.

Becker, G. S. & Murphy, K. M. (1988), 'A theory of rational addiction', *Journal of Political Economy* **96**(4), 675–700.

Cawley, J., Markowitz, S. & Tauras, J. (2004), 'Lighting up and slimming down: the effects of body weight and cigarette prices on adolescent smoking initiation', *Journal of Health Economics* **23**(2), 293–311.

Chou, S.-Y., Grossman, M. & Saffer, H. (2004), 'An economic analysis of adult obesity: results from the behavioral risk factor surveillance system', *Journal of Health Economics* **23**(3), 565–587.

Courtemanche, C. (2009), 'Rising cigarette prices and rising obesity: Coincidence or unintended consequence?', *Journal of Health Economics* **28**(4), 781–798.

Del Bono, E. & Vuri, D. (2011), 'Job mobility and the gender wage gap in Italy', *Labour Economics* **18**(1), 130–142.

Dill, D., Edwards, H. & Forbes, W. (1934), 'Tobacco smoking in relation to blood sugar, blood lactic acid and metabolism', *Am J Physiol* **109**, 118–22.

Doksum, K. (1974), 'Empirical probability plots and statistical inference for nonlinear models in the two-sample case', *The Annals of Statistics* **2**(2), 267–277.

Flegal, K. (2007), 'The effects of changes in smoking prevalence on obesity prevalence in the United States', *American Journal of Public Health* **97**, 1510–1514.

French, M. T., Norton, E. C., Fang, H. & Maclean, J. C. (2010), 'Alcohol consumption and body weight', *Health Economics* **19**(7), 814–832.

French, S. & Jeffery, R. (1995), 'Weight concerns and smoking: a literature review', *Annals of Behavioral Medicine* **17**, 234–244.

Glauser, S., Glauser, E., Reidenberg, M., Rusy, B. & Tallarida, R. (1970), 'Metabolic changes associated with the cessation of cigarette smoking', *Arch Environ Health* **20**(1), 377–81.

Gruber, J. & Frakes, M. (2006), 'Does falling smoking lead to rising obesity?', *Journal of Health Economics* **25**(2), 183–197.

Grunberg, N. (1985), 'Nicotine, cigarette smoking, and body weight', *British Journal of Addiction* **80**, 369–377.

Hiestand, W., Ramsey, H. & Hale, D. (1940), 'The effects of cigarette smoking on metabolic rate, heart rate, oxygen pulse, and breathing rate.', *J Lab Clin Med* **25**, 1013–1017.

Higgins, M.W.and Kjelsberg, M. (1967), 'Characteristics of smokers and nonsmokers in Tecumseh, Michigan', *American Journal of Epidemiology* **86**, 60–77.

Jacobs, D. & Gottenberg, S. (1981), 'Smoking and weight: The Minnesota Lipid Research Clinic', *American Journal of Public Health* (71), 391–396.

Jacobs, M., Knapp, P., Anderson, L., Karnsh, N., Meissner, R. & Richman, S. (1965), 'Relationship of oral frustration factors with heavy cigarette smoking in male college students.', *J Nerv Ment Dis* **141**, 161–171.

Karvonen, M., Keys, A., Orma, E., Fidanzo, F. & Brozek, J. (1959), 'Cigarette smoking, serum cholesterol, blood pressure and body fatness: observations in finland.', *Lancet* **1**, 492–494.

Keys, A., Aravanis, C. & Blackburn, H. (1966), 'Epidemiologic studies related to coronary heart disease: characteristics of men aged 40-59 in seven countries', *Acta Med Scand Suppl* **460**, 1–75.

Klesges, R., Meyers, A., Klesges, L. & LaVasque, M. (1989), 'Smoking, body weight, and their effects on smoking behavior: a comprehensive review of the literature', *Psychology Bulletin* **106**, 204–230.

Koenker, R. & Hallock, K. F. (2001), 'Quantile regression', *Journal of Economic Perspectives* **15**(4), 143–156.

Koenker, R. W. & Bassett, Gilbert, J. (1978), 'Regression quantiles', *Econometrica* **46**(1), 33–50.

Kowalski, A. E. (2009), Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care, NBER Working Papers 15085, National Bureau of Economic Research, Inc.

Liu, F., Zhang, N., Cheng, K.-W. & Wang, H. (2010), 'Reduced smoking and rising obesity: Does smoking ban in the workplace matter?', *Economics Letters* **108**(3), 249–252.

McCaul, K. D., Hockemeyer, J. R., Johnson, R. J., Zetocha, K., Quinlan, K. & Glasgow, R. E. (2006), 'Motivation to quit using cigarettes: A review', *Addictive Behaviours* **31**(1), 42–56.

McCormick, B. & Stone, I. (2007), 'Economic costs of obesity and the case for government intervention.', *Obesity Reviews, An Official Journal of the International Association for the Study of Obesity* **8 Suppl 1**(s1), 161–164.

Nonnemaker, J., Finkelstein, E., Engelen, M., Hoerger, T. & Farrelly, M. (2009), 'Have efforts to reduce smoking really contributed to the obesity epidemic?', *Economic Inquiry* **47**(2), 366–376.

Pieroni, L. & Salmasi, L. (2012), Does cigarette smoking affect body weight? causal estimates from the clean indoor air law discontinuity, MPRA Paper 38338, University Library of Munich, Germany.

Rashad, I., Grossman, M. & Chou, S.-Y. (2006), 'The Super Size of America: An Economic Estimation of Body Mass Index and Obesity in Adults', *Eastern Economic Journal* **32**(1), 133–148.

APPENDIX A

Table A: Descriptive statistics of covariates

	Respondents who quit smoking between 2004 and 2006 (TG_Q)				**Respondents who continued smoking between 2004 and 2006 (CG_S)**			
	2004		**2006**		**2004**		**2006**	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Male	0.49	0.5	0.49	0.5	0.46	0.5	0.46	0.5
Female	0.51	0.5	0.51	0.5	0.54	0.5	0.54	0.5
Goes for a drink at least once a week	0.37	0.48	0.29	0.46	0.34	0.47	0.33	0.47
Goes for a drink at least once a month	0.23	0.42	0.28	0.45	0.24	0.43	0.22	0.41
Goes for a drink several times a year	0.16	0.37	0.18	0.38	0.17	0.38	0.18	0.39
Goes for a drink once a year or less	0.05	0.22	0.04	0.2	0.07	0.25	0.06	0.24
Goes for a drink never/almost never	0.19	0.39	0.21	0.41	0.18	0.38	0.21	0.41
Health status is excellent	0.22	0.41	0.21	0.41	0.17	0.38	0.19	0.39
Health status is good	0.44	0.5	0.43	0.5	0.45	0.5	0.44	0.5
Health status is fair	0.23	0.42	0.23	0.42	0.25	0.43	0.24	0.43
Health status is poor	0.09	0.29	0.11	0.31	0.1	0.3	0.1	0.3
Health status is very poor	0.02	0.13	0.02	0.15	0.03	0.17	0.03	0.18
Non-long-term sick	0.94	0.24	0.95	0.22	0.91	0.28	0.91	0.28
Long-term sick	0.06	0.24	0.05	0.22	0.09	0.28	0.09	0.28
Hours worked weekly (including overtime)	39.59	11.43	38.17	12.36	37.91	13.61	37.64	12.99
Strenuous job	0.24	0.43	0.24	0.43	0.33	0.47	0.31	0.46
Non-strenuous job	0.76	0.43	0.76	0.43	0.67	0.47	0.69	0.46
Working mother	0.41	0.49	0.37	0.48	0.35	0.48	0.34	0.47
Non-working mother	0.59	0.49	0.63	0.48	0.65	0.48	0.66	0.47
Ethnic group: white	0.99	0.11	0.99	0.11	0.99	0.08	0.99	0.08
Ethnic group: other	0.01	0.11	0.01	0.11	0.01	0.08	0.01	0.08
Age 18-29	0.26	0.44	0.21	0.41	0.22	0.42	0.18	0.39
Age 30-39	0.25	0.43	0.25	0.43	0.24	0.42	0.23	0.42
Age 40-49	0.15	0.35	0.18	0.38	0.21	0.41	0.23	0.42
Age 50-59	0.12	0.33	0.12	0.33	0.18	0.38	0.18	0.39
Age 60+	0.22	0.41	0.24	0.43	0.15	0.36	0.18	0.39
1st quintile of net income	0.23	0.42	0.2	0.4	0.28	0.45	0.26	0.44
2nd quintile of net income	0.18	0.39	0.21	0.41	0.23	0.42	0.23	0.42
3rd quintile of net income	0.23	0.42	0.21	0.41	0.21	0.41	0.2	0.4
4th quintile of net income	0.2	0.4	0.2	0.4	0.16	0.36	0.17	0.38
5th quintile of net income	0.15	0.36	0.19	0.39	0.12	0.32	0.14	0.35
Couple	0.18	0.39	0.18	0.39	0.2	0.4	0.19	0.39
Married	0.46	0.5	0.5	0.5	0.43	0.5	0.44	0.5
Divorced	0.06	0.24	0.08	0.27	0.09	0.29	0.1	0.3
Separated	0.03	0.16	0.02	0.13	0.03	0.18	0.03	0.17
Widowed	0.05	0.21	0.05	0.22	0.05	0.21	0.05	0.23
Never Married	0.22	0.42	0.17	0.38	0.2	0.4	0.19	0.39
Degree	0.13	0.34	0.14	0.35	0.06	0.24	0.07	0.25
Diploma	0.26	0.44	0.29	0.45	0.25	0.43	0.28	0.45
A-level	0.11	0.31	0.1	0.3	0.12	0.32	0.11	0.31
O-level	0.17	0.38	0.17	0.37	0.19	0.39	0.18	0.38
Low educated	0.1	0.3	0.09	0.29	0.1	0.29	0.09	0.28
No qualification	0.2	0.4	0.19	0.39	0.26	0.44	0.25	0.43
Still at school	0	0	0	0	0.01	0.08	0.01	0.07
England	0.52	0.5	0.54	0.5	0.43	0.49	0.44	0.5
Wales	0.14	0.35	0.14	0.35	0.17	0.38	0.17	0.38
Scotland	0.17	0.37	0.16	0.37	0.19	0.39	0.19	0.39
Northern Ireland	0.15	0.36	0.15	0.36	0.19	0.4	0.19	0.4