The Relationship Between Obesity, Lifestyles and Employment Status

Alison Goode*
Kostas Mavromaras*
Alexandros Zangelidis+

University of Aberdeen

26 November 2004

Abstract:
This paper investigates the relationship between obesity and lifestyles (Healthy Eating and physical activity) in the light of different household income and employment status. The distinction between High and Low household income and Employed and Not-employed individuals is used in order to represent the different money and time constraints that face individuals. Typically, Healthy Eating is more expensive than less healthy lower quality food, the price of Healthy Eating usually being beyond the control of individuals. By contrast, the choice of the cost of physical activity can be controlled by individuals, but only by those who have sufficient free time. Economic theory would then suggest that where obesity may be perceived to be a problem by the individual, higher income will lead to higher consumption of (expensive) Healthy Eating and being not employed will lead to higher rate of (time consuming) physical activity. The paper starts by asking the question of whether lifestyles are endogenous when estimating BMI. Healthy Eating is found to reduce BMI and also to be endogenous. Physical activity is also found to reduce BMI, but not to be endogeneous. Instrumental variables are used in order to control for endogeneity bias and estimate individual behavioural Healthy Eating responses to higher BMI levels. The paper uses quantile regression in order to investigate these effects at different segments of the BMI distribution. Conditional estimates lend support to the predictions of economic theory: high household income individuals with high BMI respond by purchasing more Healthy Eating than their lower income counterparts. High income individuals also choose regular activities which are more expensive and less time consuming.

Acknowledgements: The provision of the data by the UK Data Archive is gratefully acknowledged. * Health Economics Research Unit; + Department of Economics. Correspondence: a.goode@abdn.ac.uk. HERU is supported by the Chief Scientist Office, Scottish Executive Health Department. The views expressed in this paper are those of the authors.

1. Introduction

The relationship between obesity and lifestyles has been well documented in the literature. Two main lifestyles feature prominently, dietary and physical activity behaviour. A large number of studies show clearly that various types of dietary and physical activity behaviour can have an impact on obesity. This paper looks at the relationship between obesity and lifestyles from an economic point of view, by introducing the dimension of household income and employment status in the picture. The main argument for introducing the economic dimension in the analysis is that both household income and employment status introduce certain constraints to individual choices. The way these constraints may impinge on the choice of lifestyles and consequently on the relationship between lifestyles and obesity can be analysed within an economic framework.

The paper presents this argument in the following steps. The first step presents a simple economic argument, whereby the time and income constraints presented to individuals by their household and employment circumstances may affect their choice of lifestyles. The second step considers the causal relationship between lifestyles and obesity by positing the question of whether lifestyles may influence obesity and obesity status may influence lifestyles. The third step brings these two strands of thought together in order to examine the degree to which a possible causal relationship between lifestyles and obesity may be influenced by time and income constraints. A brief description of these three steps follows.

First comes the economic argument. On average, individuals within the higher income households will have more money to allocate to their consumption. Part of their consumption will be related to their obesity-related lifestyles, such as Healthy Eating and physical activity. Provided that Healthy Eating and physical activity are considered to be normal goods (that is, goods that one wishes to purchase more as one gets richer), it can be expected that, other things equal, more money will be dedicated towards both Healthy Eating and physical activity by the average higher household income individual than by the average lower income individual. An analogous argument holds for those individuals who are in employment, only now the focus is on their time constraints. On average, employed individuals will have less time to dedicate to all activities outside work than their counterparts who are not employed. Hence, other things equal, the average employed individual will spend less time pursuing Healthy Eating and physical activities than the average not employed individual. In order to take such arguments on board, this paper incorporates in its modelling the possibility that income and time constraints may influence lifestyles depending on whether these lifestyles are either expensive and/or time consuming.

Second comes the causality argument. The paper recognises explicitly the possibility that there is endogeneity in the relationship between lifestyles and obesity, that is, the individual choice of lifestyles may influence obesity status *and* obesity status may influence the choice of lifestyles. Whilst there is evidence on the reducing effect of both Healthy Eating and physical activity on obesity levels measured in several ways, there is little evidence on the possible opposite effect of obesity levels on either Healthy Eating or physical activity behaviour. As a result, whilst it is well documented that Healthy Eating can reduce obesity, it is not clear whether higher obesity levels may trigger a behavioural response in terms of higher levels of Healthy Eating and/or physical activity behaviour. In order to consider the possible presence of such causal relationships, the paper builds an empirical model that allows for the testing of the presence of endogeneity and uses the technique of instrumental variables (IV) in order to correct endogeneity bias.

In order to refine the examination of the empirical relationship between obesity and lifestyles, the paper models explicitly the possibility that this relationship may vary considerably at different levels of obesity by using quantile regression. Using Body Mass Index (BMI) as a measure of obesity, the paper recognises the argument that it would be too restrictive to only look at mean effects of lifestyles on BMI. [1] Intuitively put, the paper is developed on the premise that it may be wrong to restrict estimates of the effect of any variable on BMI to be the same across the whole BMI distribution (indeed there is no reason to assume that the effect of, say, Healthy Eating behaviour will be the same for underweight and overweight individuals). Quantile regression makes no such assumption and allows semi-parametric estimation to be carried out at

[1] Of the several measures of obesity, this paper uses BMI as it is possibly the simplest and most widely used measure in the literature. Whilst the appropriateness of BMI has been questioned regarding the derivation of individual predictions, it is widely used in the context public health and large data sets analysis.

different segments of the conditional BMI distribution. This is an advantage that is particularly pertinent to the problem in hand, as it allows for the possibility that the relationship between lifestyles and BMI may be different at different levels of BMI, without imposing any theoretical priors as to what these differences may be.[2]

Using the theoretical and empirical framework described above, the paper formulates and tests the following hypotheses. Do individuals respond to higher obesity levels through more physical activity and/or Healthy Eating? If and when they do, does this response vary by the economic constraints they are faced with, represented by their household income and/or by their employed/not-employed status? Finally, do these relationships vary at different levels of BMI? In order to test these hypotheses, the paper uses the first wave of the Health and Lifestyles Survey (HALS).

The remainder of the paper is organised as follows. Section 2 presents an empirical model of BMI and lifestyles. Section 3 presents the data, Section 4 presents the estimation results and Section 5 discusses these results and Section 6 concludes. An Appendix contains detailed descriptive statistics and full estimation results.

2. The Data

2.1 The Health and Lifestyle Survey

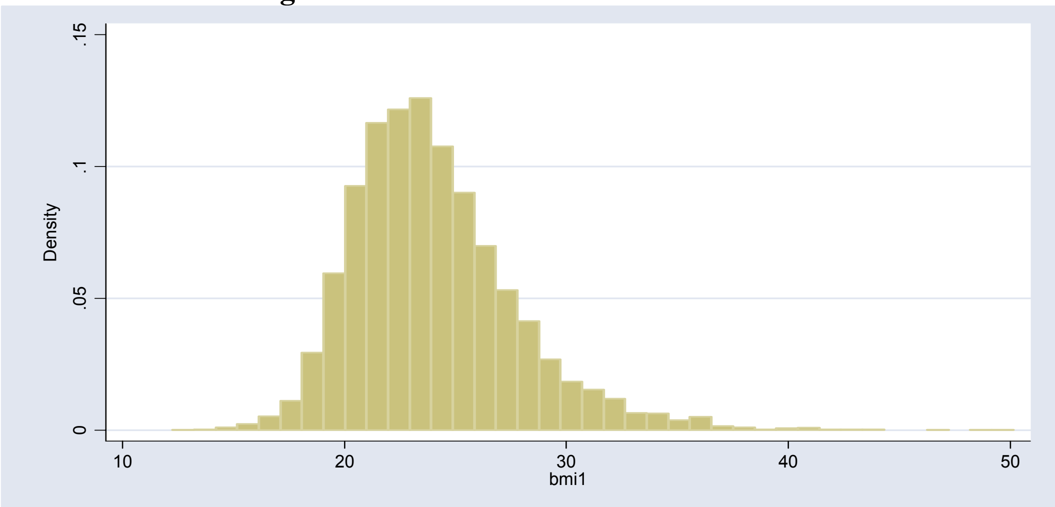
The paper uses the Health and Lifestyle Survey, which was carried out in England, Scotland and Wales during 1984 and 1985, known as HALS1.[3] The subjects were a random selection of males and females over the age of 17 living in private houses. Of the 9003 originally recorded cases, exclusion of cases over the age of 75 (489), pregnant females (113), extremely poor health cases (24), missing information on height and/or weight data (286) and missing other information (4), resulted with a sample of 8083 which has been used in this analysis. General descriptive statistics are in Appendix1.

[2] A simple way to think of quantile regression is as median regression where the data set is split into more parts than just two. Estimating median (as opposed to mean) effects removes the disproportionate effect that outliers may have. Furthermore, a more detailed split of the data, say in deciles, allows estimates to concentrate on neighbourhood rather then global effects, In a data which contains both underweight and extremely obese individuals, where reactions to BMI levels may vary considerably not only in their size, but also in their nature, there are advantages in obtaining different estimates for different segments of the conditional BMI distribution.
[3] Detailed information on the survey can be found in http://www.data-archive.ac.uk under SN: 2218.

2.2 BMI, Healthy Eating and physical activity

BMI is defined as the ratio of the weight in kilograms divided by the square of the height measured in metres. An individual who weighs 95Kgs and is 1.87m tall will have a BMI of $95/1.87^2=27.76$. BMI has been used in the analysis as a continuous variable.

Figure 1: The distribution of BMI in HALS1



Healthy Eating is based on the construction of a diet score according to the dietary quality of the food intake. This diet score takes the values of 1 (least healthy) to 30 (most healthy). The precise way in which this index is constructed is explained in Appendix 2.

Physical activity (defined as Regular Activity) is based on the information respondents provided to a question relating their own regular activity to that of other people of their own age. It takes the value 1 when someone stated that they were more active and 0 when they stated that they were as active or less active.

Table 1 shows the Healthy Eating and Regular Activity averages by Employed/Not-employed and High/Low household income.

Table 1: Healthy Eating, Regular Activity and BMI

Variable	Employed	Not-Employed	High-Income	Low-Income
BMI	23.974	24.092	23.908	24.315
	(3.506)	(4.015)	(3.419)	(3.987)
Healthy Eating	15.422	15.800	15.914	15.294
	(3.282)	(3.139)	(3.203)	(3.159)
Regular Activity	0.308	0.237	0.306	0.258
Sample size	4824	3259	2975	3503

Note: Standard deviations in brackets

3. The Empirical Model

3.1 A general model of BMI, Healthy Eating and Regular Activity

Using BMI as a measure of obesity, two lifestyles are incorporated in the analysis, Healthy Eating (*HE*) and Regular Activity (*RA*). The main equation to be estimated explains BMI in terms of lifestyles and other variables.

(1) $$BMI = \beta_1 RA + \beta_2 HE + X\gamma + u$$

where *RA* denotes Regular Activity, *HE* denotes Healthy Eating, ***X*** contains the remaining covariates of the *BMI* equation and *u* is the error term. Implicit in the analysis are two more equations, which are, however, not estimated as they are not the focus of the paper. First comes the Regular Activity equation:

(2) $$RA = \alpha_1 HE + \alpha_2 BMI + Z\delta + v$$

with *BMI* and Healthy Eating in the right hand side. ***Z*** denotes the remaining covariates of the *RA* equation and *v* is the error term. Next comes the Healthy Eating equation:

(3) $$HE = \eta_1 RA + \eta_2 BMI + Z\theta + \varepsilon$$

with *BMI* and Regular Activity in the right hand side and ε the error term. Of these three equations only equation (1) is estimated explicitly. After testing for endogeneity, equation (3) is used for the derivation of the necessary auxiliary regression for the construction of the necessary instrumental variables for equation (1). The data matrices $\boldsymbol{X}$ and $\boldsymbol{Z}$ contain some common variables, but they also contain different variables, necessary for the identification of the system.

Equation 1 is estimated using quantile regression. Quantile regression is appropriate in this instance as it assumes a common distribution of the errors across different segments of the BMI distribution, and estimates a different set of parametres for each segment of the distribution. The motivation for the use of quantile regression in the context of BMI is that a family of relationships is estimated (one for each quantile) thus allowing the examination of parameter differences across the distribution of BMI. The advantage of this method is that it allows the researcher to investigate the possibility that, not only the level of a variable may differ for different parts of the BMI distribution (for example, people who belong to different parts of the BMI distribution may have different exercise levels), but also that each "unit" of a variable may have a different effect on BMI for different parts of the BMI distribution (for example, activity may be more or less effective on people who belong to different parts of the BMI distribution).

In order to examine the possible influence of employment status on the relationship between BMI and lifestyles, the sample is split in two parts, the employed and the not employed. The same exercise is carried out for low and high household income, where the sample is split around the median income. These splits reflect one of the objectives of this paper, namely to test empirically whether the overall relationship between BMI, lifestyles and the remaining covariates in the model, may differ by employment status and/or household income. This view is confirmed by the fact that the differences in the estimates between sub-samples do not confine themselves to the lifestyles variables, but extend to many other socioeconomic and demographic variables such as, for example marital status, education and other. The implication of such differences is that the use of dummy variables and/or interaction variables in a single estimation over the complete sample would not be appropriate. Furthermore, the split of the sample between employed/not-employed and high/low income sub-samples allows the paper to test the hypothesis that the time and money constraints that face the estimated sub-samples may not only influence the choice of level of lifestyles but also the structure of the relationship between lifestyles and BMI.

3.2 Endogeneity tests and Instrumental Variables

Before proceeding with estimations the paper tests for the presence of Healthy Eating and/or Regular Activity endogeneity in Equation 1 using a Wu-Hausman test presented in Table 2 and explained in Appendix 3. Test results indicate that Healthy Eating is clearly endogenous for both employed and not employed, as well as high income and low income. By contrast, Regular Activity is found to be completely exogenous.

Table 2: Endogeneity Tests

	Healthy Eating		**Regular Activity**	
	Coefficient	*t-ratio*	*Coefficient*	*t-ratio*
Employed	.483	3.85	-.339	-0.29
F-test (d.f.: 8)	19.50		9.49	
X^2-test (d.f.: 8)	11.08			
Not Employed	.372	2.45	1.02	0.57
F-test (d.f.: 8)	14.18		5.07	
X^2-test (d.f.: 8)	7.48			
Employed	.470	4.93	-	-
F-test (d.f.: 8)	18.55			
X^2-test (d.f.: 8)	11.08			
Not Employed	.423	3.22	-	-
F-test (d.f.: 8)	14.18			
X^2-test (d.f.: 8)	6.50			
High-Income	.525	3.67	-1.628	-1.33
F-test (d.f.: 8)	13.38		7.32	
X^2-test (d.f.: 8)	9.49			
Low-Income	.355	1.61	-.076	-0.03
F-test (d.f.: 6)	11.73		4.77	
X^2-test (d.f.: 6)	5.60			
High-Income	.428	3.66	-	-
F-test (d.f.: 8)	12.52			
X^2-test (d.f.: 8)	10.68			
Low-Income	.357	2.19	-	-
F-test (d.f.: 6)	11.29			
X^2-test (d.f.: 6)	5.76			

Note: Numbers reported here are the coefficients (and their robust t-ratios) of the estimated residuals of the Healthy Eating and the Regular Activity variables on BMI. The F-test refers to the strength of the instrumenting variables in the estimation which generated the (predicted) instrument. The value of a Staiger Chi square test is also reported. STATA was used for all estimations.

The IV method implemented here amounts to replacing the endogenous variables with their predicted values in order to re-write Equation 1. Predicted values are obtained from the following auxiliary equation:

(4) $HE = \eta_1 RA + Z\eta_2 + \varepsilon$

where ***Z*** contains *all* exogenous variables in the model, including the ones used to identify the system. Predicted values $H\hat{E}$ are then used to estimate

(5) $BMI = \beta_1 RA + \beta_2 H\hat{E} + X\gamma + u$

By construction, the estimates derived from Equation 5 and reported in Tables 3 and 4 in the next section and in Appendix 4 are free of endogeneity bias. One important aspect of this type of analysis is the choice of variables used to generate the predicted values which are used as the instrument $H\hat{E}$ for Healthy Eating variable in Equation 5. These variables (which are included in *Z* in estimation 4 and not in *X* in estimation 5) have to exhibit the following statistical properties. First, they must be correlated with the variable that needs to be instrumented (here, Healthy Eating).[4] Second, they should not be correlated with the error term u in Equation 5. Intuitively put they may not have a direct impact on the explained variable (here, BMI).[5] These tests are reported in Table 2 below each coefficient and t-ratio.

4. Quantile regression results

Having established that only Healthy Eating can be shown to be endogenous, Equation 5 was estimated using quantile regression and using IV only for Healthy Eating. That is, in the right hand side of Equation 5 Healthy Eating was entered in the form of its prediction derived from Equation 4 and Regular Activity was assigned its observed values. Given the large number of tables generated by quantile regression (nine tables for each sub-sample) summary results showing the sign and the statistical

[4] This can be tested formally using an F-test of their joint strength in Equation 4.
[5] This can be tested by regressing them on the explained variable and deriving the Staiger statistic which equals the (number of observations minus the degrees of freedom) times R^2. This statistic follows the χ^2 distribution.

significance of the estimates by household income category and by employment status are presented in Tables 3 and 4 respectively.[6]

Table 3 shows rather clearly that Healthy Eating, where it is shown to be significantly different from zero, has a negative effect on BMI. There are strong signs of this negative effect for high household income individuals (Table 3, column 1) which appears from the third decile and increases with BMI. By contrast, the low household income sub-sample shows only scatterred and weak signs of significance in deciles 6, 7 and 8 (Table 3, column 3). Regular Activity is shown to have a negative effect on BMI for both income categories. The BMI of low household income individuals appears to be strongly negatively affected across the whole BMI distribution (Table 3, column 4). The BMI of high household income individuals is negatively affected, but only for those above median BMI (Table 3, column 2). It is clear that the empirical patterns which arise from Table 3 show considerable differences between the low and high household income groups. The implications of these differences are discussed in detail in the next section.

Table 3: Statistical Significance of Healthy Eating and Regular Activity Effect on BMI by Household Income Category

	High Household Income		Low Household Income	
BMI Decile	1. Healthy Eating	2. Regular Activity	3. Healthy Eating	4. Regular Activity
1	o	o	o	--
2	o	o	o	---
3	-	o	o	---
4	-	o	o	--
5	--	---	o	--
6	--	o	--	--
7	---	--	-	--
8	---	--	-	--
9	---	---	o	--

Note: 0: statistical significance less than 10%; -, --, ---: significant at 10%, 56% and 1% respectively.

Table 4 shows results of the sample split between the employed and the not-employed. Like with the high/low household sample split, both Healthy Eating and Regular Activity have a negative effect on BMI where they show any statistical significance. Unlike with the low/high household income sample split, estimates in

[6] Full results are contained in an extended Appendix available from the authors by request.

Table 4 suggest that the effects of Healthy Eating are rather similar between the employed and the not employed sub-samples (columns 1 and 3), with only deciles 3 and 4 that showing a difference.

Table 4: Statistical Significance of Healthy Eating and Regular Activity Effect on BMI by Employment Status Category

	Employed		Not Employed	
BMI Decile	1. Healthy Eating	2. Regular Activity	3. Healthy Eating	4. Regular Activity
1	o	o	o	---
2	o	-	o	---
3	--	o	o	--
4	--	--	o	---
5	---	o	--	--
6	---	o	---	---
7	---	--	---	---
8	---	---	---	---
9	---	---	---	o

Note: 0: statistical significance less than 10%; -, --, ---: significant at 10%, 56% and 1% respectively.

The effect of some of the other covariates in the regressions on BMI are worth noting. The reader should be reminded that these are partial effects, i.e. they are over and above the effects of all other variables included in the regression. Alcohol intake does not appear to have a major effect on BMI. It is interesting to note that a positive effect of being a regular drinker on BMI is estimated, but only amongst the lower segments of the BMI distribution and only for the high household and employed sub-samples. By contrast, being a regular smoker has a strong negative effect on BMI for all sub-samples and across the whole of the BMI distribution. Education estimates suggest very clearly that higher education levels have a negative impact on BMI which manifests itself in part through healthier eating. Being married has a positive effect on BMI, but only below median BMI and only for the high household income and the employed sub-sample. Above median BMI no statistically significant effect can be traced. In a similar fashion being employed (entered as a dummy variable in the high/low household income sub-samples) appears to influence BMI positively, but only below median BMI.[7]

[7] Regional and age dummies were also used in the regressions as control variables. Their detailed effects can be seen in the detailed Appendix.

5. Discussion

The section concentrates on the interpretation of the results of the paper. Interpretation revolves around two crucial parts of the analysis. First, the ability of quantile regression to estimate effects at different parts of the conditional BMI distribution.[8] Second, on the ability of this analysis to distinguish the causal direction of effects. This is done by contrasting estimations which do not use the IV method, which reflect the overall association between BMI and Healthy Eating, and estimations which use the IV method, which reflect the *net* effect of Healthy Eating on BMI (net in the sense that they have controlled for the reverse endogenous effect of BMI on Healthy Eating). By taking the difference between these two estimates one obtains an indirect estimate of the effect of BMI on Healthy Eating. These points are best explained by way of example. Figure 2 presents such a comparison for the high household income sub-sample.

Figure 2:BMI and Healthy Eating (High Household Income)

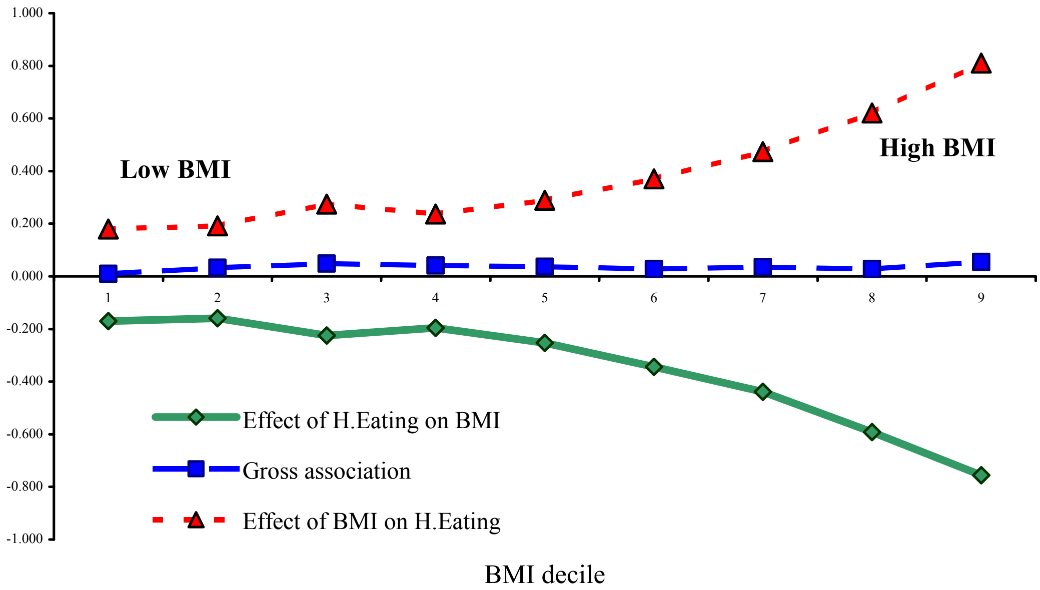


Figure 2 is based on the results of two estimations. To begin with, the gross association between BMI and Healthy Eating was estimated without instrumenting for the endogeneity of Healthy Eating. However, given that endogeity tests have

[8] These should be distinguished from results depending on the unconditional (observed) BMI distribution. The remainder of this section refers to the conditional distribution.

confirmed that the relationship between BMI and Healthy Eating is a two-way relationship, the gross association estimates are biased in the sense that they contain both the effect of Healthy Eating on BMI and the effect of BMI on Healthy Eating. The line marked “effect of Healthy Eating on BMI” plots the results from the second estimation, the IV quantile regression. The line marked “effect of BMI on Healthy Eating” is the difference between the gross association and the effect of Healthy Eating on BMI. This line is an indirect estimate of the behavioural response of individuals in terms of their different Healthy Eating because of their BMI status.

A first inspection of Figure 2 would show that the gross association between BMI and Healthy Eating is close to zero and not statistically significant for all BMI levels. As soon as the information contained in the IV regression is considered the picture changes. The “effect of Healthy Eating on BMI” IV estimates suggest that there is empirical support for the existence of endogenous Healthy Eating behaviour and that this is principally based on what happens at the higher levels of the BMI distribution. It is only for the high household income sub-sample and at the higher BMI levels that an estimated behavioural response appears in the form of the upward sloping “effect of BMI on Healthy Eating” line. The result that individuals respond to their BMI level by adjusting their dietary behaviour makes good intuitive sense, as one would expect that a behavioural Healthy Eating response would only be triggered at BMI levels which are considered as “too high” by the individuals in question.[9] Similar quantile regression estimates reveal that the low household income sub-sample fails to produce statistically significant results.[10]

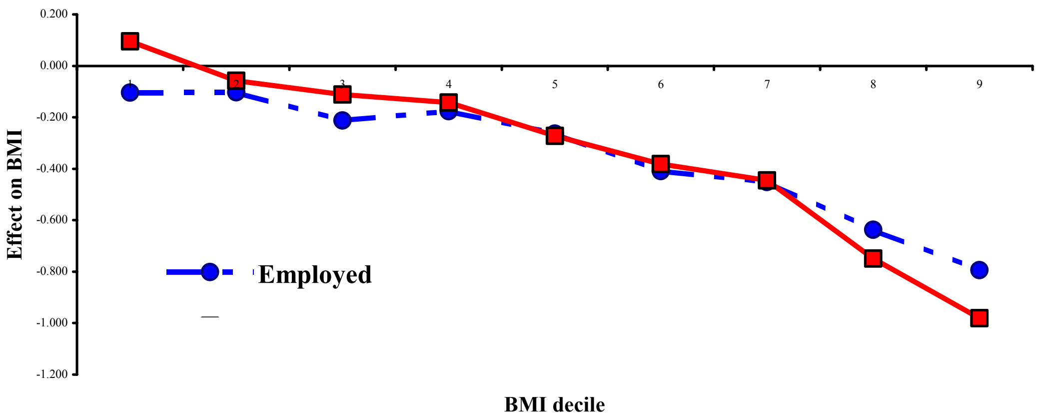
Results suggest that, when it comes to Healthy Eating choices, household income seems to matter. Individuals in high income households respond to higher levels of BMI through higher levels of Healthy Eating. This response starts at about the median of the BMI distribution and increases with BMI. Low household income individuals do not generate statistically significant results regarding Healthy Eating. The

[9] If there is behaviour which is of a preventive nature, that is lifestyles which result in preventing obesity, the data in hand cannot trace it. It is only the response to higher BMI that is estimated by this analysis. In this sense the estimates presented here could be considered as under-estimates of obesity related behaviour.

[10] In the low household income sub-sample there is no support for the hypothesis that the estimated effects are statistically significantly different from zero. It is noteworthy, however, that the size of the not significant low income sub-sample coefficients is very similar to the size of the significant high income sub-sample coefficients.

implication is that the money constraint seems to matter regarding Healthy Eating choices. Two possible interpretations are offered for this result. First, it could be that both high and low income households have similar preference structures, but only the high income households can afford to respond to individual obesity through more expensive Healthy Eating. Second, it could be that the data is not sufficiently informative and there is some unobserved difference in the preference structure which influences Healthy Eating choices. To the degree that the multivariate nature of the data and the estimation control for preference differences (through variables such as age, educations and other), the former interpretation should be chosen, namely, that low household income individuals would want to respond to higher BMI levels, but given their general low level of consumption capabilities, consumption on Healthy Eating is also lower. The possibility of individual unobserved preference differences which vary systematically by household income is one which cannot be tested with the data in hand and will remain an open question in this analysis.

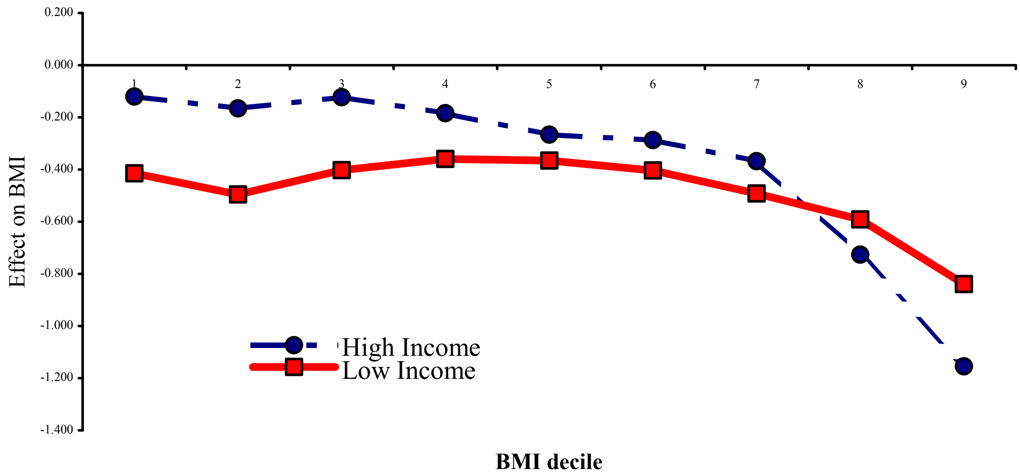
Figure 3: Effect of Healthy Eating on BMI by Employment Status



Estimations splitting the sample into the employed and the not-employed produces very similar results for both sub-samples, suggesting that, when it comes to Healthy Eating, employment status matters little. Healthy Eating starts to matter from around median BMI with an increasing effect as BMI increases, the effect being independent of employment status. Given that the main clear-cut difference between the employed and the not-employed is their time constraint, this result suggests that there is no empirical support for the proposition that time constraints may matter regarding Healthy Eating choices.

Unlike Healthy Eating, there is no empirical support for the endogeneity of Regular Activity. That is, the data suggests that Regular Activity levels influence BMI levels, but BMI levels do not influence Regular Activity levels. The effect of Regular Activity on BMI is always negative, suggesting that those who exercise benefit from a lower BMI. The responsiveness of BMI to Regular Activity at low BMI levels is of little interest within the context of this research. Results on the responsiveness of BMI to Regular Activity at higher BMI level are of interest. The investigation of the upper deciles of the BMI distribution in Figure 4 suggests that Regular Activity is more effective amongst individuals in higher income households.

Figure 4: Effect of Regular Activity on BMI by Household Income



It should be remembered at this stage that Regular Activity refers to the perceived comparison of individuals with comparators of their own age. As such, differences in the effectiveness of Regular Activity in reducing BMI by household income would make little sense. A possible explanation of this, apparently counter-intuitive, estimate may be that the type of activity differs in a systematic way by household income level (see Table 6), whereby intensive (and more expensive) activities form a larger part of the higher household income individuals portfolio of activities. Table 6 shows rather clearly this to be the case, with high household individuals reporting a much higher proportion of sports and physical leisure activities amongst their overall activities (a total of 45.9% of high household income individuals report that they "play particular sports" and/or they have other "Physical leisure activities") than their low income counterparts (with a total of 27.3% of the same activities). By contrast, Table 5 shows low household income individuals to take part in lower intensity activities such as walking, gardening and housework (total of 39%) more than their high income counterparts (20.8%).

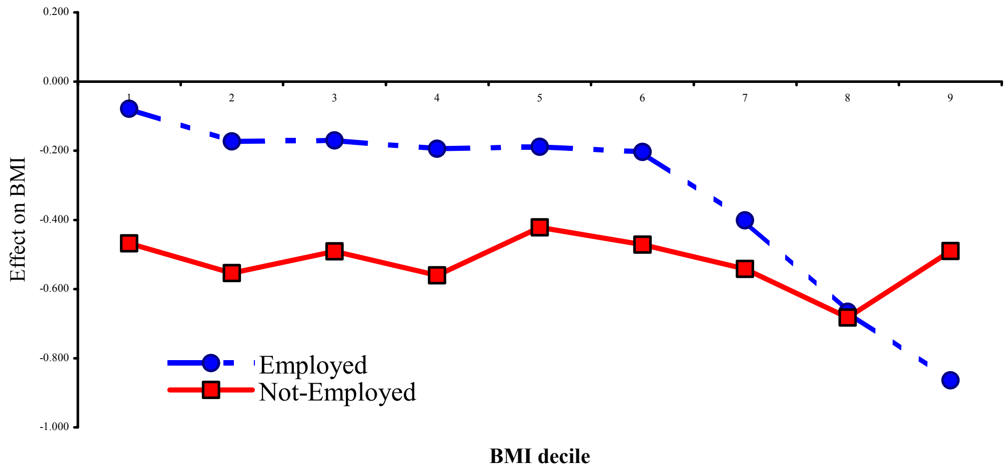
Table 5: Types of activities by household income

	High Household Income	Low Household Income
"Anything to keep healthy"	64.84	60.38
"Housework"	1.11	3.43
"Gardening"	5.14	7.74
"Walking"	17.24	22.58
"Play particular sports"	22.05	10.42
"Physical leisure activities"	23.87	16.90

Note: Percentage of people doing particular activities by household income level.

A similar, but not as clear, empirical pattern arises in Figure 5, which makes the distinction between employed and not employed individuals.

Figure 5: Effect of Regular Activity on BMI by Employment Status



Although the BMI of the employed appears to be less sensitive to Regular Activity, it appears that the employed engage more in less time consuming activities, such as sports and physical leisure, and the not-employed engage more in more time consuming activities such as gardening, housework and walking.

Table 6: Types of activities by employment status

	Employed	Not Employed
"Anything to keep healthy"	62.83	62.66
"Housework"	0.83	4.82
"Gardening"	4.33	9.14
"Walking"	15.57	25.04
"Play particular sports"	20.92	9.76
"Physical leisure activities"	22.20	18.26

Figures 4 and 5 and Tables 5 and 6 confirm the possibility that Regular Activity was reported in this data set using individuals in similar income/employment status as comparators. This suggestion goes some way towards explaining the differences in the estimated effectiveness of Regular Activity at reducing BMI reported in Figures 4 and 5. Furthermore, they provide some indirect evidence that time and budget constraints play a role in the choice of lifestyles.

6. Conclusion

This paper investigated the relationship between obesity and lifestyles in the light of different household income and employment status. The distinction between high and low household income and employed and not-employed individuals was used in order to represent the different money and time constraints that face individuals. The paper first asked whether the lifestyles of Healthy Eating and Regular Activity are endogenous when estimating BMI, that is, whether lifestyles may influence obesity status *and* obesity status may influence lifestyles. Healthy Eating was found to be endogenous, regular exercise was not.

The paper introduced the method of instrumental variables in order to control for the endogeneity of Healthy Eating in the BMI equation. The paper introduced quantile regression in order to estimate the relationship between lifestyles and BMI across the BMI distribution. Results suggested that Healthy Eating influences obesity and obesity influences Healthy Eating but only at the higher levels of the BMI distribution and for the high income individuals. This suggested that income constraints influence Healthy Eating patterns. When the sample was split into the employed and the not-employed (where time constraints are different), results suggested that, in terms of

Healthy Eating, employment status matters little. These two results suggest that, as a response to high BMI levels, Healthy Eating is subject to money but not time constraints.

Unlike Healthy Eating, Regular Activity was not found to be endogenous. The effect of Regular Activity was always negative and results suggested it was more effective among individuals in higher income household. A similar result was found for the employed. This may be explained by the fact that both those in the employed and high income categories participate in different types of activities than those individuals in the not-employed and low income categories. That is confirmed by the fact that both the employed and high income individuals report a much higher proportion of sports and more intense physical leisure activities (which are more expensive and less time intensive) and a much lower proportion of low intensity activities (which are less expensive and more time consuming) than do low income and not-employed individuals.

References

Arias, O., K.F. Hallock, and W. Sosa-Escudero 2001, "Individual Heterogeneity in the Returns to Schooling: Instrumental Variables Quantile Regression Using Twin Data", *Empirical Economics Journal*, 26, 7—40.

Buchinsky, M. 1998, "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research", *Journal of Human Resources*, 33, 88-126.

Davidson, R, and J. G. MacKinnon 1993, *Estimation and Inference in Econometrics*, Oxford University Press, New York.

Goode, A 2003 "Associations of Body Mass Index with diet, physical activity, alcohol consumption and smoking: The Grampian Adult Lifestyle Survey 1998", MSc thesis, Department of Public Health, University of Aberdeen, Scotland, UK.

Goode, A and K Mavromaras 2004 "Estimation of the Effect of Diet on BMI for Males and Females using Instrumental Variables", mimeo, University of Aberdeen.

Goode, A, K Mavromaras and A Zangelidis 2004 "Investigating the relationship between BMI and Smoking, Alcohol Intake, Exercise and Healthy Diet using Quantile Regression" mimeo, University of Aberdeen.

Guitiérrez-Fisac JL, Rodríguez Artalejo F, Cuallar-Castillon P, Banegas Banegas JR, del Rey Calero J. 1999 "Determinants of geographical variations in body mass index (BMI) and obesity in Spain." *International Journal of Obesity; 23(4): 342-347.*

Hakala P, Rissanen A, Koskenvuo M, Kapiro J, Rönnemaa T. 1999 "Environmental factors in the development of obesity in identical twins." *International Journal of Obesity; 23 (7): 746-753*

Health Promotions. Grampian Adult Lifestyle Survey. 1998 Report.

Laaksonen M, Prattala, R, Helasoja V, Uutela A, Lahelma E. 2003 "Income and health behaviours. Evidence from monitoring surveys among Finnish adults. *Journal of Epidemiology & Community Health. 57(9):711-717.*

Lahti-Koski M. Pietinen P. Heliovaara M. Vartiainen E. 2002 "Associations of body mass index and obesity with physical activity, food choices, alcohol intake, and smoking in the 1982-1997 FINRISK Studies." *American Journal of Clinical Nutrition; 75(5): 809-17.*

Lahti-Koski M. Vartiainen E. Mannisto S. Pietinen P. 2002 "Age, education and occupation as determinants of trends in body mass index in Finland from 1982 to 1997." *International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 24(12):1669-76.

Mannisto S, Pietinen P, Haukka J, Ovaskainen M-L, Albanes D, Virtamo J. 1996 "Reported alcohol intake, diet and body mass index in male smokers." *European Journal of Clinical Nutrition; 50(4): 239-245.*

Rissanen AM, Heliovaara M, Knekt P, Reunanen A, Aromaa A. 1991 "Determinants of weight gain and overweight in adult Finns." *European Journal of Clinical Nutrition; 45(9): 419-30*

Sherwood NE, Jeffery RW, French SA, Hannan PJ, Murray DM. 2000 "Predictors of weight gain in the Pound of Prevention Study." *International Journal of Obesity; 24(4): 395-403.*

Slattery NE, McDonald A, Bild DE, Caan BJ, Hilner JE, Jacobs Jr DR, Liu K. 1992 "Associations of body fat and its distribution with dietary intake, physical activity, alcohol, and smoking in blacks and whites." *American Journal of Clinical Nutrition; 55(5): 943-9.*

Staiger D and J H Stock 1997 "Instrumental Variables Regression with Weak Instruments" *Econometrica; 65: 557-586.*

Tavani A, Negri E, Vecchia C. 1994 "Determinants of body mass index; a study from Northern Italy." *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity; 18(7):497-502.*

Wooldridge J M. 2002 "Econometric Analysis of Cross Section and Panel Data", The MIT Press, London, England.

Appendix 1: Descriptives

	High Income				Low Income			
Variable	*Mean*	*St. Dev.*	*Min.*	*Max.*	*Mean*	*St. Dev.*	*Min.*	*Max.*
BMI	23.91	3.42	12.25	43.88	24.32	3.99	14.27	50.14
HE hat	15.91	1.56	11.13	19.94	15.52	1.46	10.87	20.02
HE	15.91	3.20	3	29	15.29	3.16	3	26
Reg. Active	0.31		0	1	0.26		0	1
Reg. drinker	0.33		0	1	0.20		0	1
Reg. smoker	0.30		0	1	0.39		0	1
Male	0.51		0	1	0.43		0	1
Married	0.87		0	1	0.67		0	1
Age17-24	0.10		0	1	0.09		0	1
Age25-34	0.25		0	1	0.19		0	1
Age35-44	0.31		0	1	0.15		0	1
Age45-54	0.20		0	1	0.14		0	1
Gen. Health	1.98	0.72	1	4	2.24	0.82	1	4
Wales	0.05		0	1	0.07		0	1
North	0.04		0	1	0.08		0	1
NorthWest	0.12		0	1	0.13		0	1
York-Humb	0.07		0	1	0.10		0	1
Midlands	0.16		0	1	0.18		0	1
London	0.12		0	1	0.08		0	1
Scotland	0.07		0	1	0.12		0	1
EducSec	0.27		0	1	0.18		0	1
EducProf	0.33		0	1	0.17		0	1
EducUni	0.09		0	1	0.02		0	1
Employed	0.81		0	1	0.41		0	1
Attitude1	3.75	0.80	1	5	3.93	0.83	1	5
Attitude2	1.76	0.83	1	5	1.61	0.75	1	5
Attitude3	3.25	0.93	1	5	3.13	0.99	1	5
Behaviour	0.31		0	1	0.21		0	1
SpouseWork	0.69		0	1	0.27		0	1
SEGProfess	0.10		0	1	0.02		0	1
SEGManage	0.24		0	1	0.11		0	1
SEGNonMan	0.23		0	1	0.17		0	1
Sample size	2975				3503			

Appendix 1: Descriptives (cont)

	Employed				Not-employed			
Variable	*Mean*	*St. Dev.*	*Min.*	*Max.*	*Mean*	*St. Dev.*	*Min.*	*Max.*
BMI	23.97	3.51	15.15	46.55	24.09	4.01	12.25	50.14
HE hat	15.45	1.56	10.57	19.94	16.01	1.55	11.12	20.38
HE	15.42	3.28	3	29	15.80	3.14	2	29
Reg. Active	0.31	0.46	0	1	0.24	0.43	0	1
Reg. drinker	0.32	0.47	0	1	0.17	0.37	0	1
Reg. smoker	0.33	0.47	0	1	0.35	0.48	0	1
Male	0.53	0.50	0	1	0.32	0.47	0	1
Married	0.74	0.44	0	1	0.70	0.46	0	1
Age17-24	0.15	0.36	0	1	0.10	0.30	0	1
Age25-34	0.23	0.42	0	1	0.15	0.36	0	1
Age35-44	0.27	0.45	0	1	0.12	0.32	0	1
Age45-54	0.20	0.40	0	1	0.11	0.31	0	1
Gen. Health	2.02	0.73	1	4	2.27	0.84	1	4
Wales	0.06	0.23	0	1	0.05	0.23	0	1
North	0.05	0.21	0	1	0.08	0.27	0	1
NorthWest	0.12	0.32	0	1	0.13	0.34	0	1
York-Humb	0.09	0.28	0	1	0.09	0.29	0	1
Midlands	0.16	0.37	0	1	0.18	0.38	0	1
London	0.11	0.32	0	1	0.09	0.29	0	1
Scotland	0.10	0.30	0	1	0.11	0.32	0	1
EducSec	0.26	0.44	0	1	0.20	0.40	0	1
EducProf	0.29	0.45	0	1	0.17	0.38	0	1
EducUni	0.07	0.25	0	1	0.02	0.14	0	1
Employed	1.00	0.00	1	1	0.00	0.00	0	0
Attitude1	3.78	0.81	1	5	3.94	0.84	1	5
Attitude2	1.74	0.82	1	5	1.59	0.75	1	5
Attitude3	3.23	0.96	1	5	3.13	0.97	1	5
Behaviour	0.29	0.45	0	1	0.20	0.40	0	1
SpouseWork	0.55	0.50	0	1	0.31	0.46	0	1
SEGProfess	0.06	0.23	0	1	0.04	0.20	0	1
SEGManage	0.19	0.39	0	1	0.16	0.36	0	1
SEGNonMan	0.23	0.42	0	1	0.17	0.37	0	1
Sample size	4824				3259			

Appendix 2: The calculation of Healthy Eating variable

The diet score was created using a scoring system developed by Lanarkshire Health Board, Scotland, UK. Ten foods were selected, five healthy foods (fruit, vegetables, rice/pasta, potatoes and pulses) and five less healthy foods (chips, other fried foods, cakes, crisps and sausages/meat filled pies). Respondents were awarded a score in accordance with how often per week they consumed each of the ten foods, 0 = least healthy, 3=most healthy. These individual scores were then added together resulting in a final score ranging between 0 and 30 where 30 represented the healthiest diet.

Appendix 3: Testing for the presence of endogeneity

Take a model of two equations with dependent variables *BMI* (Equation 1) and *Healthy Eating* (Equation 2). Suppose that variable *Healthy Eating* may be suspected of endogeneity when it is included in the right hand side of the *BMI* Equation 1. The test for endogeneity used in this paper is run as follows. First, all exogenous variables of the model (that is, all variables that are exogenous and appear in the right hand side of either Equation 1 or Equation 2 are regressed on variable *Healthy Eating* which is suspected of endogeneity. The estimated residuals from this regression are retained. Second, the estimated residuals from Equation 2 are included in the regression of *BMI* alongside with observed *Healthy Eating*. This is Equation 5 in the main text. The null hypothesis is that there is no endogeneity. If the estimated coefficient of the residuals is significantly different from zero, the null hypothesis can be rejected. In practical terms, the rejection of the null would suggest that the relationship between *BMI* and *Healthy Eating* is a two-way relationship where *Healthy Eating* causes *BMI* to change and *BMI* causes *Healthy Eating* to change. For an exposition of this type of Hausman test see Davidson and MacKinnon (1993). It is important to note that this test is always subject to the model specification that is used. In other words, if the specification is not right (possibly in the sense of some crucial variables missing) then the test may provide no evidence of endogeneity when there actually is endogeneity. Simply put, the rejection of the null can be trusted as evidence of the existence of endogeneity. The failure of the test to reject the null, however, can be due to either mis-specification or to lack of endogeneity. It is, therefore, important that a large number of specifications are tried out in order to assess the robustness of the model. The results presented in this paper have undergone extensive such robustness tests.

Appendix 4: Quantile regression estimates

BMI Quantile Estimates									
Decile	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
High Income									
Healthy Eating	-0.170	-0.159	-0.225	-0.195	-0.252	-0.343	-0.439	-0.592	-0.755
	(-1.39)	(-1.41)	(-1.87)	(-1.85)	(-1.98)	(-2.42)	(-3.57)	(-3.42)	(-2.86)
Regular Activity	-0.121	-0.165	-0.123	-0.184	-0.267	-0.287	-0.368	-0.727	-1.155
	(-0.73)	(-1.12)	(-0.78)	(-1.36)	(-1.64)	(-1.59)	(-2.34)	(-3.36)	(-3.42)
Low Income									
Healthy Eating	0.274	-0.047	-0.225	-0.214	-0.191	-0.367	-0.398	-0.459	-0.557
	(1.50)	(-0.32)	(-1.44)	(-1.47)	(-1.03)	(-1.94)	(-1.68)	(-1.69)	(-1.35)
Regular Activity	-0.415	-0.495	-0.402	-0.360	-0.366	-0.404	-0.492	-0.591	-0.839
	(-2.14)	(-3.31)	(-2.53)	(-2.42)	(-1.93)	(-2.11)	(-2.08)	(-2.21)	(-2.12)

Note: Sample sizes, high income (2975), low income (3503). Dependent variable: continuous BMI. Quantiles indicate cut off points for the estimation, (e.g. quantile 1 was estimated leaving 10 percent to the left and 90 percent to the right). t-ratios in parentheses.