Obesity and health care costs: Some overweight considerations

Albert Kuo, Ted Lee, Querida Qiu, Geoffrey Wang

May 14, 2015

Abstract

This paper investigates obesity's impact on annual medical expenditures using a two-part model (2PM). The two-part model is estimated using publicly available data from the Medical Expenditure Panel Survey for 2000–2012. With our results, we confirmed earlier studies that obesity has a significant effect on annual medical costs, especially for women and the elderly. However, in view of shorter life expectancies for obese individuals, we further investigated whether obesity has a net cost or gain effect on medical costs. We concluded that while obesity still has a net cost, the cost is smaller, particularly for women, suggesting that from a public policy standpoint, the costs of obesity may be overstated.

1 Introduction

Since the 1970s, obesity has been a rapidly growing problem in the United States and around the world [1, 2, 3]. There is a consensus on the greater costs, both economic and otherwise, associated with the obesity epidemic. Studies have shown that obesity leads to a decrease in life expectancy and healthy life-years [4]. Moreover, with the use of novel methods such as instrumental variables (IV), it has been suggested that the impact of obesity on medical costs has long been understated [5].

While the use of an IV model in [5] does make up for many of the limitations seen in earlier literature in estimating the impact of obesity on medical costs, we have identified a few improvements that can be made to the model. First, we now have access to a wider time frame of data from the medical (2000-2012). Second, we use a more precise definition of obesity, categorizing obesity into four levels: healthy, overweight, obese, and very obese. Third, we use the group lasso method to confirm that among the variables we had available, family weight is the best choice as an IV. Finally, we estimate a two-part model (2PM) using a Box Cox transformation instead of a gamma generalized linear model, which allows us to choose an optimal transformation to obtain normally distributed residuals.

In addition, the conclusion in the study drawn by [5] is limited by its failure to account for the shorter life expectancy associated with obesity. Other studies have sought to investigate the association between obesity, life expectancy, and medical expenditure over the course of a lifetime, such as that conducted on the Ohsaki Cohort Study [6]. As a parallel, there has been much debate within the literature of the net effects of smoking on healthcare costs, given smokers' shorter life expectancies. In this paper, we similarly investigate the net effects on healthcare costs due to obesity by considering lifetime medical expenditures.

2 Data

We use the Medical Expenditure Panel Survey (MEPS) data from the year 2000 to 2012. Like Cawley and Meyerhoefer, we also limit our data to adults between the ages of 20 and 64. We then use consumer price index (CPI) data from the Bureau of Labor Statistics (BLS) to correct for inflation by converting all expenditures into 2012 dollars.

For our analysis of lifetime medical expenditures, we use life tables and life years lost due to obesity in the study by [4], who used data from 2003 to 2010 from the National Health and Nutrition Examination Survey (NHANES).

3 Empirical Model

3.1 Exploratory Analysis

We estimate linear models, setting body mass index (BMI)[1] as our primary explanatory variable and the total annual expenditure on health care as our response variable. We reconstruct the results of [5] by fitting a linear model for data from years 2000 to 2005. Then we improve upon their estimates by using data from years 2000 to 2012 to estimate the same model. We also conducted linear regression on subgroups in the data, divided by sex, age, and obesity levels [2].

3.2 Choice of instrumental variable

[5] is the first paper to use the instrumental variable (IV) approach, in order to correct for two shortcomings of past research: endogeneity of weight on healthcare expenditure and errors in weight data stemming from self-reported surveys.

We noted that [5] did not externally validate their choice of the IV. To select our IV, we used a graphical model to model the dependencies among all the scale and interval variables (i.e., age, total income, total expenditure, total paid by private institution, average family weight, BMI, and number of months since last blood check). For estimation of the graphical model, we assumed the model was multivariate Gaussian and estimated the precision matrix. To do this, we first log-transformed them into a Gaussian distribution. Then, we used the `glasso` [7] for the optimization, which has the advantage of tracking which between-variable dependency has a greater signal. Intuitively, the strongest relationships come in first.

At penalty parameter $\lambda = 3.012$, `glasso` detected the edge signal of BMI and average family weight. This suggested that this is an ideal choice as an

[1] Body mass index is defined as weight in kilograms divided by height in meters squared.

[2] Obesity levels are healthy (BMI between 18.5 and 25), overweight (BMI between 25 and 30), obese (BMI between 30 and 35), and very obese (BMI greater than 35).

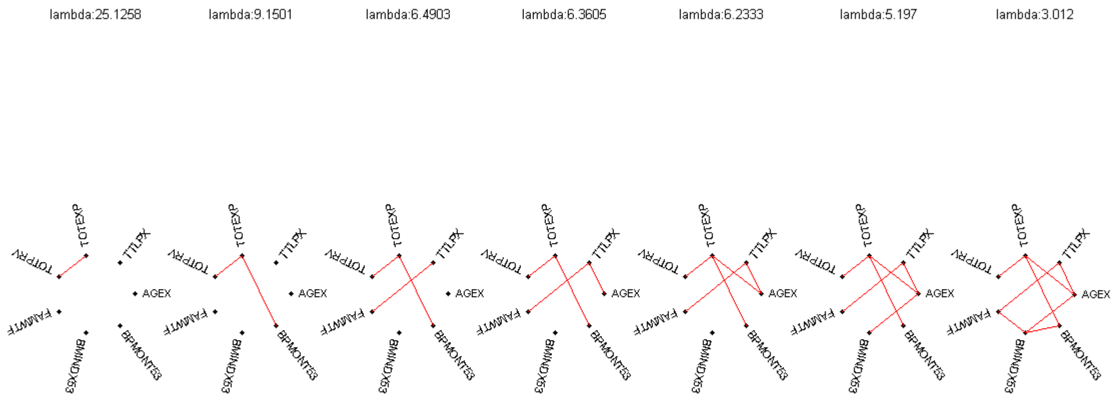

Figure 1: *Estimated conditional dependencies (markov field) according to varying regularization parameter (λ). It is apparent that Average Family Weight is the first (and in fact, only) node to be connected with* `BMINDX` *(BMI) before being connected to* `TOTEXP` *(total expenditure).*

instrument, as its dependence with BMI is more prominent than total expenditure. Hence, we decided to use the average family weight as IV. Our IV, average family weight, is similar to the IV used by [5] (BMI of oldest offspring).

3.3 2PM

Since many respondents have zero medical expenditures, we use a 2PM consisting of a logit model in the first part and a Box-Cox transformed ordinary linear regression in the second part to model skewed data.

The first part defines a dichotomous variable d, defined by

$$d = \begin{cases} 1 & \text{if } y^* = \mathbf{x}^\top \boldsymbol{\tau} + \epsilon_1 > 0.5 \\ 0 & \text{if } y^* \leq 0.5, \end{cases}$$

where y^* is the latent variable (probability of having nonzero medical expenditures), $\mathbf{x}$ is the covariate (BMI), $\boldsymbol{\tau}$ is the estimated coefficient, and ϵ_1 is the error term, assumed to follow $N(0, \sigma)$. Our $\boldsymbol{\tau}$ can be computed using a logit estimation [9].

The second part uses ordinary linear regression for the restricted case where $d = 1$, i.e.,

$$\mathbb{E}[y|y > 0, \mathbf{x}] = \mathbf{x}^\top \beta + \mathbb{E}[\epsilon_2 | y > 0, \mathbf{x}].$$

where y is the dependent variable (total annual medical expenditure), and ϵ_2 is the error term. Using 2PM, we obtain estimates for the impact of obesity on total medical expenditures that are less likely to be biased.

3.4 Box-Cox transformation

[2] and [5] use a generalized linear model with a log-link and gamma distribution for the second part, justified by the right-skewness of the medical expenditures. Nonetheless, their simple log-transformation may not be the optimal transformation to fulfill the assumption of normality, since there is no guarantee that the response variable is lognormal.

For our linear regression in 2PM, we use another method, i.e. a Box-Cox transformation, to obtain residuals that are closer to a normal distribution. A plot of the parameter values for Box-Cox indicated that the parameter 0.2 is optimal, which leads to a transformation of $y^{\text{Box}-\text{Cox}} = y^{0.2}$. This is different from the log transformation used by [5].

3.5 Lifetime Medical Expenditures

To estimate lifetime medical expenditure, M_i, we use the existing results from [4] on the life expectancies and the life years lost due to obesity. Due to differences in life expectancies among age groups and between males and females, we calculated separate lifetime medical expenditures for each group. In general, the lifetime medical expenditure is calculated as follows:

$$M_i = (A + A_i) * (L - L_i)$$

where A is our estimated total annual medical expenditure for healthy individuals, A_i is the increase in total annual medical expenditure for obesity level i, L is the expected life expectancy for healthy individuals and L_i is the number of life years lost for obesity level i.

The lifetime medical expenditure increase (LMEI) is the difference between the lifetime medical expenditure for the obese sample and the healthy sample and is given by

$$\text{LMEI}_i = M_i - A * L$$

.

To compare our results with our estimated annual medical expenditure in 2PM, we also calculated the average annual lifetime medical expenditure. Average annual lifetime medical expenditure, $\bar{M_i}$ is calculated by dividing lifetime medical expenditures by the life expectancy, i.e.

$$\bar{M_i} = \frac{M_i}{L - L_i}$$

The adjusted annual medical expenditure increase (AMEI) is the difference between the annual lifetime medical expenditure for obese individuals and healthy individuals and is given by

$$\text{AMEI}_i = \bar{M_i} - A$$

4 Results

4.1 Exploratory Analysis

Table 1 shows our simple linear regression results: as expected, BMI is positively correlated with total medical expenditures. We ran a simple linear regression model on both the 2000 to 2005 data and the 2000 to 2012 data. The lower coefficient for 2000 to 2012 data indicates that with more years of data included in the model, the average increase in total healthcare expenditures with a unit increase in BMI is about $14 less than the average increase for the 2000 to 2005 data. However, in both cases, the coefficient is significant.

Table 1: BMI on Healthcare Expenditure, by years

	Dependent variable:	
	Total Expenditures	
	2000 to 2005	2000 to 2012
BMI	121.446***	107.148***
	(3.908)	(5.549)
Constant	35.654	327.901**
	(111.841)	(157.056)
Observations	227,213	94,604
Note:	*p<0.1; **p<0.05; ***p<0.01	

For the rest of our analyses, we only use the more comprehensive data set from 2000 to 2012.

Table 2: BMI on Healthcare Expenditures, by Sex

	Dependent variable:	
	Total Expenditure	
	Male	Female
BMI	114.695***	125.685***
	(7.401)	(4.245)
Constant	−373.949*	449.744***
	(210.682)	(122.056)
Observations	106,983	120,230
Note:	*p<0.1; **p<0.05; ***p<0.01	

We examine obesity's impact on medical costs on the subgroups of men and women (Table 2). For both men and women, BMI has a significantly positive effect on total annual expenditures, with a slightly higher impact for women, which is consistent with Cawley and Meyerhoefer. Our regression shows that a unit increase in BMI results in an increase of \$115 in medical expenditures on average for men and \$126 on average for women.

Table 3: Weight Category on Expenditure, by Age Group

	Total Expenditure			
	Age			
	20 to 34	35 to 44	45 to 54	55 to 64
Overweight	-186.053^{**}	-433.629^{***}	140.290	540.462^{***}
	(91.869)	(103.001)	(124.426)	(194.337)
Obese	122.541	57.001	915.534^{***}	$1{,}858.859^{***}$
	(117.386)	(122.416)	(145.281)	(225.064)
Very Obese	737.207^{***}	$1{,}420.873^{***}$	$2{,}765.310^{***}$	$4{,}206.602^{***}$
	(138.382)	(140.298)	(164.514)	(255.933)
Constant	$1{,}967.210^{***}$	$2{,}694.339^{***}$	$3{,}508.330^{***}$	$5{,}164.899^{***}$
	(61.172)	(76.003)	(92.926)	(147.717)
Observations	80,995	54,502	52,750	38,966
Note:			*p<0.1; **p<0.05; ***p<0.01	

We then use dummy variables for obesity level based on BMI [3].

From Table 3, we observe that in general, being either overweight, obese, or very obese increases annual medical expenditures; being very obese has the most significant effect, with an increase of up to \$4,207 for respondents between the ages of 55 to 64. Moreover, we can see clearly that the impact of obesity on medical care expenditures increases with age across the board.

4.2 2PM and Box-Cox

For the first part of the two part model, we run a logit model to estimate the probability of positive medical expenditures. From Table 4, we see

$$\log(\frac{p}{1-p}) = 0.317 + .015(\text{BMI})$$

Therefore, when BMI = 0, there is a 31.7% probability of positive expenditure. In addition, for a unit increase in BMI, the relative odds of positive medical expenditures increases by 1.5%, which is significant at the 1% level.

Using a Box-Cox transform, we use $\lambda = 0.2$ which is similar to the gamma GLM with a log-link approach, but has a theoretical guarantee of the residuals being distributed normally. For the second part of the two part model, we ran an adjusted[4] linear regression, with the transformed dependent variable. Our

[3]Obesity levels are healthy (BMI between 18.5 and 25), overweight (BMI between 25 and 30), obese (BMI between 30 and 35), and very obese (BMI greater than 35).

[4]We adjusted for other covariates, including race, education, and marriage status

Table 4: Part 1: Logistic Regression

	Dependent variable:
	P(Positive Total expenditure)
BMI	0.015***
	(0.005)
Constant	0.317
	(0.014)
Observations	227,213
Note:	*p<0.1; **p<0.05; ***p<0.01

results in Table 5 show that total medical care expenditures increase on average by $93.28 for each unit increase in BMI. In addition, it shows that a one year increase in age will, on average, increase expenditures by $126.628, so the older a patient gets the more that he will spend on healthcare. Finally, we see that females are expected to spend more on healthcare ($1081.16 more than men).

Table 5: Part 2: Box-Cox Transform Adjusted Linear Regression

	Dependent variable:
	Total expenditure
BMI	93.279***
	(3.981)
Age	126.628***
	(2.287)
Female	1,081.158***
	(49.668)
Constant	-5,211.611***
	(0.380)
Observations	175,231
Note:	*p<0.1; **p<0.05; ***p<0.01

4.3 Lifetime Medical Expenditures

We further investigate lifetime medical expenditures due to obesity, which better reflects the implications of obesity for public health policies and its broader societal impact. Although annual medical costs increase with obesity on average, obesity also leads to shorter life expectancies [4]. Thus, it is not immediately apparent how obesity affects the total medical expenditures over one's lifetime.

Table 6: Lifetime Medical Expenditures Increase for Men

Age group	20-39	40-59	60-79
Overweight	-2612.86	4186.00	720.21
Obese	19075.28	20147.12	10377.16
Very Obese	82100.31	60133.15	36130.46

Table 7: Lifetime Medical Expenditures Increase for Women

Age group	20-39	40-59	60-79
Overweight	-160.41	5.86	-44.66
Obese	642.69	528.43	251.76
Very Obese	3591.95	1962.44	1447.13

The lifetime annual medical expenditure increase (LMEI) for men and women are provided in Table 6 and 7, respectively. We see that in general, LMEI is positive and much greater for men. This is unexpected, since without taking life expectancy into account, it appears that obesity's impact on medical costs is greater on women (see Table 2). Although for most obese individuals, LMEI is positive, for overweight individuals between the ages of 20 and 39, there appears to be a decrease in lifetime medical expenditures. Lastly, for older individuals, the LMEI is lower regardless of sex or obesity level, which can be explained by the fewer number of years they have left. Thus, the number of life years they lose due to obesity is smaller.

Table 8: Average Annual Medical Expenditures Increase for Men

Age group	20-39	40-59	60-79
Overweight	-92.65	148.44	25.54
Obese	676.43	714.44	367.98
Very Obese	2911.36	2132.38	1281.22

Table 9: Average Annual Medical Expenditures Increase for Women

Age group	20-39	40-59	60-79
Overweight	-5.69	0.21	-1.58
Obese	22.79	18.74	8.93
Very Obese	127.37	69.59	51.32

In comparison to Table 3, the average annual medical expenditure increase (AMEI), as described by Tables 8 and 9, we see that after taking life expectancies into account, over the course of a lifetime, the increase in medical costs due to obesity is less in all cases. While, we do not have same age groups between two data sets, we can still see a clear trend that Table 3 presents a bigger impact of obesity on total expenditures than those presented by Tables 8 and 9.

5 Discussion

From our initial regression results, a higher BMI, i.e. obesity, leads to a significant increase in annual medical expenditures. Such impact is slightly higher for women than men and also higher for older age groups. Our logit model also confirms that a higher BMI has a significant impact on increasing the probability of incurring positive medical care costs, confirming the results of [5].

[5] showed that weighing an additional unit of BMI is associated with $49 higher annual expenditures from their pooled sample. Our results indicate that obesity healthcare costs are actually greater that the ones predicted in [5], as shown in Table 5.

First, we see that using a wider range of years in the data for our analysis, the effect on total medical expenditure with a unit increase in BMI is smaller. Next, using a 2PM model with a logit as the first part and a Box-Cox transformation for the second part, we get similar results as [5]: total healthcare expenditure increases with each unit increase in BMI. In addition, we see a positive relationship between age and medical expenditure and females pay more for healthcare compared to males, corroborating the claim of [5] that there are large differences between males and females.

For lifetime medical expenditures, similar problems have been studied quite extensively for smokers, with mixed results. For example, according to [13] (2008), among a group of healthy-living individuals, obese individuals, and smokers in Netherlands, lifetime health expenditure was highest among healthy-living individuals and lowest for smokers. But according to [12], lifetime medical costs of male heavy smokers are 47% higher than for men who never smoked. The contrast in findings might be due to a large difference in the amount of medical care used by smokers relative to nonsmokers in the United States and Netherlands data sets, among other model differences. Nonetheless, this topic bears significant weight for smoking-related public health policies. Similarly, it would be revealing to study the economic impact of obesity over lifetime. Due to time and data constraints, we were not able to perform a complete regression from the original dataset. Instead, as discussed earlier, we used the parameters estimated and reported by [4], and used the formulas specified in Section 3.5 to estimate lifetime medical expenditure increase. Our results for lifetime medical expenditures show that the net effects of obesity on healthcare costs are less than previously estimated. After adjusting for shorter life expectancies, we found that the AMEI is greatest for young men who are very obese, with an increase of up to $2911. However, the effects are much smaller for individuals who are overweight or obese and for older individuals. Moreover, the net effect for obese women is substantially reduced when taking into account their shorter life expectancies, with young women who are very obese having an increase of up to \$127. This suggests that especially for women, the net impact of obesity on healthcare costs from a public policy perspective is not as dramatic after taking life expectancies into account. This is in contrast to the results of [5], where obesity was shown to lead to an substantial increase in annual medical expenditures.

A difference from previous research is that while [5] only compared obese to non-obese (both healthy weight and overweight), we separated our sample population into five categories: healthy, normal/ideal, overweight, obese, and very obese and compared the medical expenditures of each obese group with the expenditures of the healthy group.

Further research would test the significance of the difference in lifetime medical expenditure between healthy and obese individuals after taking into account shorter life expectancies associated with obesity. In our study, we were able to give some evidence that a higher BMI is correlated with a significant increase in annual medical expenditures for older men based on past studies done of healthcare costs taking into consideration life expectancy. However, due to time constraints, we were unable to obtain confidence intervals for the increase in average lifetime medical expenditure after adjusting for shorter life expectancies.

Additional improvements to our empirical model could be made by investigating different models that correct for skewness in data. Due to time constraints, we largely followed the 2PM used by [5], albeit using a transformation as determined by the Box-Cox method, but the effectiveness of different models could also be tested in further analysis.

References

[1] Finkelstein, E.A., Fiebelkorn, I.C., Wang, G., 2003. National medical spending attributable to overweight and obesity: how much and who's paying? Health Affairs, W3–W219.

[2] Finkelstein, E.A., Trogdon, J.G., Cohen, J.W., Dietz, W., 2009. Annual medical spending attributable to obesity: payer- and service-specific estimates. Health Affairs, Web Exclusive. July 27, 2009.

[3] Finkelstein, E.A., Zuckerman, L.L., 2008. The Fattening of America. John Wiley & Sons,Inc., NJ.

[4] Grover et al., 2015. Years of life lost and healthy life-years lost from diabetes and cardiovascular disease in overweight and obese individuals: a modelling study.

[5] Cawley, John and Meyerhoefer, Chad, The medical care costs of obesity: An instrumental variables approach, Journal of Health Economics 31 (2012) 219–230.

[6] Nagai, Masato, et al. "Effect of age on the association between body mass index and all-cause mortality: the Ohsaki cohort study." Journal of epidemiology 20.5 (2010): 398-407.

[7] Friedman, Jerome, Trevor Hastie, and Rob Tibshirani. "glasso: Graphical lasso-estimation of Gaussian graphical models." R package version 1 (2008).

[8] Bound, J., Brown, C., Mathiowetz, N., 2002. Measurement error in survey data. In:Heckman, J., Leamer, Ed. (Eds.), Handbook of Econometrics, vol. 5. Springer-Verlag, New York, pp. 3705–3843.

[9] Humphreys, Brad R. "Dealing with zeros in economic data." (2010).

[10] Buntin, Melinda Beeuwkes and Zaslavskyb, Alan M., Too much ado about two-part models and transformation?: Comparing methods of modeling Medicare expenditures, Journal of Health Economics 23 (2004) 525-542.

[11] Fontaine, K.R., Bartlett, S.J., 2000. Access and use of medical care among obese per-sons. Obesity Research 8, 403–406.

[12] Hodgson, Thomas A. "Cigarette smoking and lifetime medical expenditures." The Milbank Quarterly (1992): 81-125.

[13] Van Baal, Pieter HM, et al. "Lifetime medical costs of obesity: prevention no cure for increasing health expenditure." PLoS medicine 5.2 (2008): e29.