Changes in Obesity Prevalence in Canada

Ye Tao
PhD Candidate in Economics
Concordia University
Supervisor: Nikolay Gospodinov
Co-supervisor: Ian Irvine

Motivation

- As many developed countries, obesity is becoming a very severe social problem in Canada, the prevalence of obesity among adults aged 18 or older has been increasing significantly since the end of 1970s', from 13.8% in 1978/79 to 23% in 2004 (Tjepkema, 2006).

- Almost all studies on the impact of obesity reveal strong positive relationship between the excess weight and diseases such as cardiovascular disease, diabetes and hypertension. Obesity is also a reason of increased prevalence of psychological disorders, such as depression (T.A.Wadden et al., 2002). As a person's weight jumps to a high level of obesity, the risk of having these diseases increases dramatically (Allison et al., 1999; Engeland et al., 2003; Flegal et al., 2005).

- Many researchers and medical professionals believe that obesity will overtake smoking as the most important cause of morbidity and premature mortality in the near future if the current trend of the obesity epidemic continues.

Motivation

- Study on obesity shows that once people gain weight, it is hard to get rid of it. Hence, the most effective way to reduce the future impact of obesity to the health and health care system is preventing the occurrence of obesity.
- Undoubtedly, a detailed factor analysis of obesity and reliable estimation of the influence by these factors could help not only the government but also the individuals.
- Few papers on comparing factors that contribute to obesity and even fewer papers on the evolution of the effects of these factors over time.

Brief introduction of this paper

- This study focuses on the prevalence of obesity in Canada. By applying the quantile regression and backward elimination method to three health surveys, this paper attempts to detect two kinds of changes :

  1. Changes of BMI that are attributable to factors at different points of the BMI distribution.

  2. Changes of the importance of these factors over time.

- Based on these analyses, a prediction of the future prevalence of obesity in Canada is projected.

Data

--- Survey selection

- Three health surveys
  1. Canada Health Survey conducted in 1978/1979
  2. National Population Health Survey (Cycle 1) conducted in 1994/1995
  3. Canadian Community Health Survey (Cycle 2.2, Nutrition) conducted in 2004.

  are selected for this study and all these surveys were conducted by Statistics Canada.

- Seven factors were selected to reflect people's body mass index in this research. They are resident region, age, marital status, education, working status, family income and physical activity index.

- Each factor has several dummy variables to represent different groups or characteristics of people.

Data

--Variables

Table 1: Definition of the Variables

Factors	Covariates	Notation
Region	Atlantic Provinces	reg^1
	Quebec	reg^2
	Ontario	reg^3
	Prairie Provinces	reg^4
Age	35 to 49	age_g^2
	50 to 64	age_g^3
Marital Status	Married	$mars^2$
	Widowed/Separated/Divorced	$mars^3$
Education	Secondary to Some Post Secondary	ed^2
	Post Secondary	ed^3
Working Status	Currently Working	ws^2
	Not Currently Working	ws^3
Family Income	Lower Middle Level	inc_q^2
	Middle Level	inc_q^3
	Upper Middle Level	inc_q^4
	Highest Level	inc_q^5
Physical Activities	Moderate	pai^2
	Active	pai^3

Model

The quantile regression model

$$bmi_i = X_i^{'}\beta_\theta + \varepsilon_{\theta i} \quad \text{with} \quad Quant_\theta(bmi_i \mid X_i) = X_i^{'}\beta_\theta \quad (i = 1, ..., n) \tag{1}$$

where β_θ and X_i are $k \times 1$ vectors, and $x_{i1} \equiv 1$. $Quant_\theta(bmi_i \mid X_i)$ denote the θth conditional quantile of body mass index given X.

While the corresponding linear regression model is

$$bmi_i = X_i^{'}\beta + \varepsilon_i \quad (i = 1, ..., n) \tag{2}$$

Estimation Result

---Quantile regression VS Linear regression

The estimation of the quantile regression and Linear regression shows that:

the estimation from the linear regression is significantly different from estimation from the quantile regression for most covariates at most of the quantiles, particularly at the high quantiles, which mostly represent people with high BMI.

Linear regression only reflects the relationship between BMI and the factors at average level; however, we are much interested in people in the higher levels of BMI. Hence, the following results only reflect the results from the quantile regression estimation.

Estimation Results

--- Comparison of Percentage distribution of BMI

In order to check the accuracy of estimation of this study, the percentage distribution of BMI from the quantile regression was compared with estimation reported by Statistics Canada (Tjepkema 2006) for 1978 and 2004.

Table 8: Percentage Distribution of BMI

Year	Sex	Under weight	Normal	Overweight	Obese I	Obese II	Obese III
1978	Male	0	48%	40.7%	10.3%	1%	%
	Female	4.1%	56.9%	28.7%	8.4%	1.9%	%
1978*	Male	1.3	44.6%	42.5%	9.5%	F	F
	Female	3.4%	52%	28.7%	11.5%	2.9%	1.5%
2004	Male	0	32.4%	41.1%	18.6%	6.5%	1.4%
	Female	2%	43.5%	30.6%	13.6%	5.8%	4.5%
2004*	Male	1.4	33.6%	42%	16.5%	4.8%	1.6%
	Female	2.5%	44.1%	30.2%	14%	5.5%	3.8%

The data of 1978* and 2004* are adopted from Table 2 of Tjepkema (2006)

Estimation

--- Estimated Distribution of BMI

If we take a look of the estimated conditional distribution from the quantile regression for years of 1978 1994 and 2004, more information will be found.

Figure 1

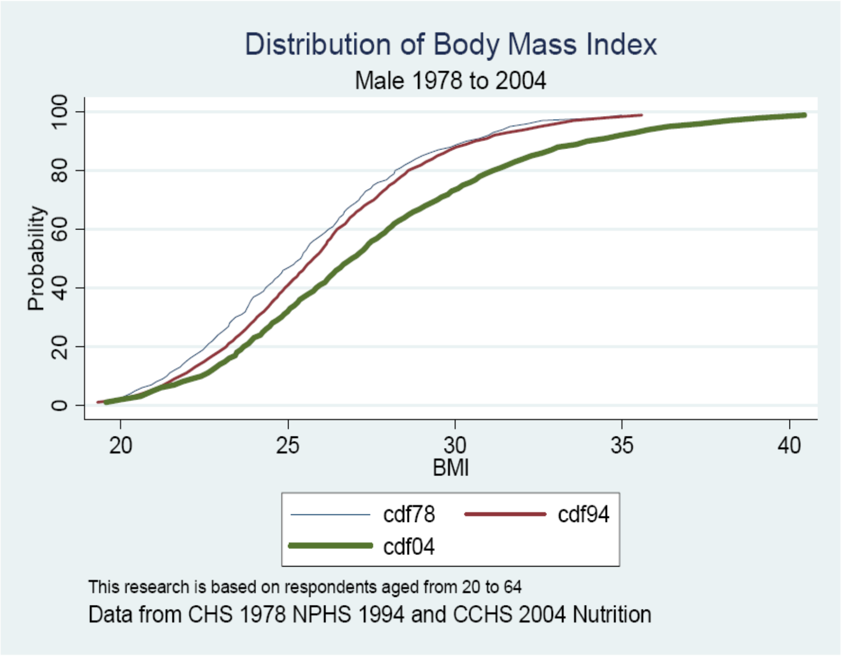
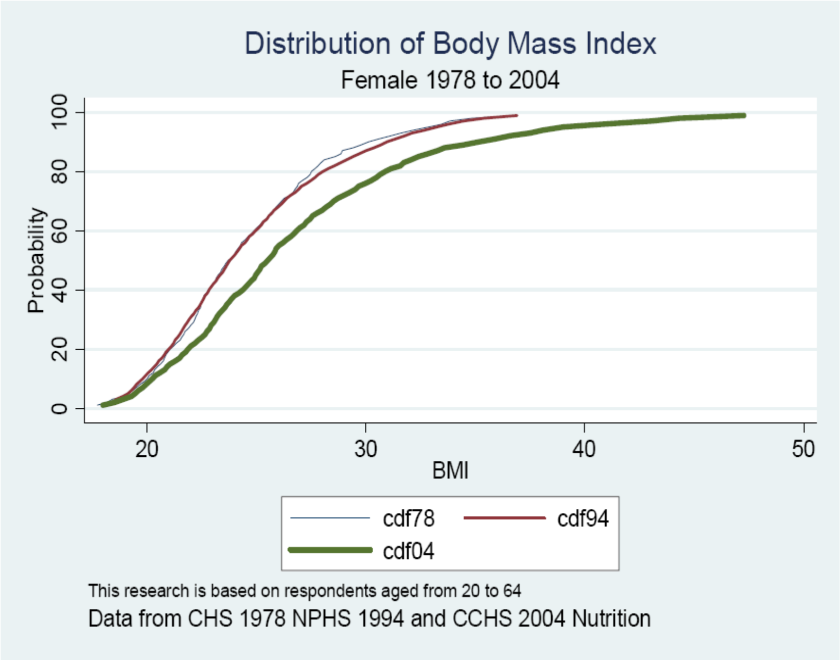


Figure 2



The importance of the factors

--- Backward elimination

To identify factors which have largest impact on BMI, A method called backward elimination is used in this research.

- Backward elimination is a method which uses the amount of unique variance a variable adds to the complete model (all remaining variables) as the criterion for exclusion from the model.

- In this method, the full model (with all variables included) is computed first. Then each variable is removed from the model alternatively, and the variable that causes the least reduction in accounted variance by its removal from the model is the first to be eliminated.

- This technique was applied to the selected factors for the model in this paper.

The importance of the factors---Men

Table 9: Order of Importance of Factors (Male)

Quantile	1	2	3	4	5	6	7
	CHS1978						
10	Age	Income quantile	Physical activity	Region	Marital status	Working status	Education
25	Age	Region	Marital status	Physical activity	Income quantile	Working status	Education
50	Age	Region	Marital status	Income quantile	Physical activity	Working status	Education
75	Marital status	Income quantile	Physical activity	Age	Region	Working status	Education
90	Income quantile	Working status	Marital status	Physical activity	Age	Region	Education
95	Income quantile	Working status	Physical activity	Marital status	Age	Region	Education
99	Physical activity	Region	Income quantile	Working status	Age	Education	Marital status
	NPHS1994/95						
10	Age	Region	Marital status	Working status	Income quantile	Education	Physical activity
25	Age	Marital status	Region	Education	Income quantile	Working status	Physical activity
50	Age	Marital status	Education	Region	Income quantile	Working status	Physical activity
75	Education	Marital status	Region	Age	Physical activity	Income quantile	Working status
90	Education	Region	Age	Income quantile	Physical activity	Working status	Marital status
95	Education	Region	Age	Physical activity	Marital status	Income quantile	Working status
99	Education	Region	Working status	Age	Physical activity	Income quantile	Marital status
	CCHS2004 Nutrition						
10	Marital status	Income quantile	Age	Education	Region	Working status	Physical activity
25	Age	Income quantile	Marital status	Physical activity	Region	Working status	Education
50	Marital status	Region	Education	Income quantile	Age	Working status	Physical activity
75	Age	Physical activity	Income quantile	Working status	Education	Region	Marital status
90	Education	Physical activity	Income quantile	Region	Age	Marital status	Working status
95	Education	Physical activity	Region	Age	Income quantile	Working status	Marital status
99	Region	Education	Physical activity	Income quantile	Age	Marital status	Working status

The importance of the factors---Women

Table 10: Order of Importance of Factors (Female)

	CHS1978						
Quantile	1	2	3	4	5	6	7
10	Age	Region	Working status	Physical activity	Income quantile	Marital status	Education
25	Age	Region	Working status	Physical activity	Income quantile	Marital status	Education
50	Age	Income quantile	Region	Physical activity	Working status	Marital status	Education
75	Age	Working status	Region	Income quantile	Physical activity	Marital status	Education
90	Age	Income quantile	Physical activity	Region	Working status	Marital status	Education
95	Age	Income quantile	Physical activity	Working status	Region	Marital status	Education
99	Physical activity	Region	Age	Marital status	Working status	Income quantile	Education
	NPHS1994/95						
10	Age	Region	Marital status	Working status	Income quantile	Education	Physical activity
25	Age	Region	Marital status	Working status	Income quantile	Physical activity	Education
50	Age	Region	Income quantile	Marital status	Physical activity	Education	Working status
75	Age	Region	Working status	Physical activity	Income quantile	Education	Marital status
90	Age	Income quantile	Region	Physical activity	Working status	Marital status	Education
95	Income quantile	Physical activity	Region	Age	Marital status	Working status	Education
99	Physical activity	Income quantile	Region	Education	Marital status	Working status	Age
	CCHS2004 Nutrition						
10	Age	Region	Marital status	Income quantile	Education	Physical activity	Working status
25	Age	Region	Marital status	Education	Income quantile	Physical activity	Working status
50	Age	Education	Physical activity	Region	Income quantile	Marital status	Working status
75	Physical activity	Age	Education	Region	Income quantile	Marital status	Working status
90	Physical activity	Region	Age	Education	Income quantile	Marital status	Working status
95	Physical activity	Education	Region	Income quantile	Age	Marital status	Working status
99	Physical activity	Working status	Income quantile	Region	Age	Marital status	Education

The importance of the factors

- The age, region and marital status are always important factors for both men and women's body mass index, especially at the low quantiles.
- At high quantiles, the influence of marital status to the body mass index becomes pretty weak. Although the influence of region and age to the body mass index also become weak, in general, their influences are still strong over time.
- Family income is very important factor for overweight and obese women in 1978 and 1994. In 2004, it is not an important factor except women with highest BMI.
- For overweight and obese people, the importance of education and physical activities increases over the period from 1978 to 2004; specifically, at the high quantiles, education plays the most important role in men's BMI and physical activities plays the most important role in women's BMI.

Prediction in BMI

--- Decomposition of the difference of BMI distribution

To predict future BMI distribution, the difference of BMI distribution between different surveys at each quantile from 1st to 99th quantile was decomposed. The BMI distributions which were used for the decomposition are the estimations obtained with the quantile regression. The difference of BMI at θ th quantile can be denoted as follow.

$$bmi_\theta^T - bmi_\theta^t = \bar{X}^{T'}\beta_\theta^T - \bar{X}^{t'}\beta_\theta^t = \left(\bar{X}^{T'} - \bar{X}^{t'}\right)\beta_\theta^t + \bar{X}^{t'}\left(\beta_\theta^T - \beta_\theta^t\right) + \left(\bar{X}^{T'} - \bar{X}^{t'}\right)\left(\beta_\theta^T - \beta_\theta^t\right) \quad (4)$$

The pure influence caused by the change of structure of the factors: $\left(\bar{X}^{T'} - \bar{X}^{t'}\right)\beta_\theta^t$

The pure influence caused by the change of function of the factors: $\bar{X}^{t'}\left(\beta_\theta^T - \beta_\theta^t\right)$

Intersectional influence: $\left(\bar{X}^{T'} - \bar{X}^{t'}\right)\left(\beta_\theta^T - \beta_\theta^t\right)$

Prediction in BMI

--- The results of the decomposition

By comparing the three components of the decomposition formula for the difference of the BMI distribution, this study find that:

- The influence caused by the change of function of the factors is the dominant component of the difference of BMI distribution between 1994/95 and 2004 for most quantiles.

- There is no dominant component in the difference of BMI distribution between 1978 and 1994/95.

.

Prediction in BMI

--- Two implication for prediction in BMI distribution

Combining the previous result (slide 10), which is changes in BMI distribution is significant over 1994/95 to 2004 period and is quite insignificant from 1978 to 1994/95, the results of decomposition have two very important implications to the prediction of body bass index:

- Significant shift in BMI distribution is caused by the change of function of the factors.
- The change in BMI distribution caused by the change of structure of the factors is insignificant.

Prediction in BMI

--- Discussion of possible way for prediction

If we interpret the prediction of prevalence of obesity as a **warning** rather than a reliable forecast, significant shift in BMI distribution deserves more attention. Therefore, Based on the above two implications and the formula for decomposition (slide 15), the prediction for BMI at year T could be expressed as

$$bmi_{\theta}^{T} = bmi_{\theta}^{t} + \bar{X}^{t'}\left(\beta_{\theta}^{T} - \beta_{\theta}^{t}\right) + bias$$

Prediction in BMI

--- Discussion of possible way for prediction

If the prediction is based on the survey in 2004, then the prediction is

$$bmi_\theta^T = bmi_\theta^{04} + \bar{X}^{04'}\left(\beta_\theta^T - \beta_\theta^{04}\right) + adjustment$$

If we also assume that

- the tendency of BMI evolution in the next decade from 2005 is same as the decade that ended at 2004, and second,
- the change of BMI occurs evenly overtime.

$\beta_\theta^T - \beta_\theta^{04}$ can be estimated as $\left(\beta_\theta^{04} - \beta_\theta^{94}\right) * (T - 2004)/9.5$ where 9.5 represents the length of time from the middle of the survey NPHS94/95 to survey CCHS2004.

And adjustment could be expressed as the product of $\bar{X}^{04'}\left(\beta_\theta^T - \beta_\theta^{04}\right)$ and average quotients that defined in different sections of distribution of BMI.

Prediction in BMI

--- Results of the prediction

By using the above formula for prediction, the predicted percentage distributions of BMI in 2007 and 2014 are given in Table 11.

Table 11: Predicted Percentage Distributions of BMI

Sex	Male			Female		
Year	2007^A	2007^P	2014^P	2007^A	2007^P	2014^P
Under weight	1.2%	0	0	3.9%	1.1%	1.8%
Normal	38.4%	28.8%	23.5%	50.2%	41%	36%
Overweight	39.3%	40.5%	38.8%	25.7%	31.8%	30.5%
Obese I	13.3%	20.8%	21.8%	9.8%	15.7%	17.4%
Obese II	2.7%	7.2%	10.5%	3.4%	5.9%	6.5%
Obese III	1.2%	2.7%	5.4%	1.8%	5.6%	9.6%

A denote estimation from Statistics Canada (CANSIM, Table 105-4009) and P denote the predicted percentage distribution of BMI from this study. Our prediction is for people aged 20 and 64, while estimation from Statistics Canada is for people aged 18 and over.

Thanks