An Analysis of the Effect of Income on Adult Body Weight in China

Ruizhi Zhang
Faculty Advisor: Professor David Roland-Holst

University of California, Berkeley
Department of Economics
May 2015

Abstract: Even though there have been a plethora of studies and theories in literature about the positive correlation between income and body weight in developing countries, there are few studies on body weight trends over the years, especially regional trends where economic development differs. This paper uses the case of China and studies body weight growth trends as average income rises for regions with different levels of economic development. The analysis finds out that adult body weight experiences a three-stage growth—the overcoming hunger phase, the leveling phase, and the affluence phase. In addition, while the current relationship between income and body weight for Chinese adults is still positive, and overweight and obesity problems are more prevalent in more developed regions, future body weight growth resides in the less developed areas.

Acknowledgement

I would like to express my sincere gratitude to my advisor, Professor David Roland-Holst, who has contributed time and invaluable advice to guide me through every step of the development of this paper, and encouraged me when obstacles arose during the research process. I would also like to thank my friend, Hui Ding, for her help during my data cleaning process. Last but not least, I would like to thank my parents—my most cherished and respected ones, for their continuous support and love.

I. Introduction

Overweight and obesity induced by excessive body fat has been studied for decades for its prevalence and solutions. An increasing body weight not only hinders worker productivity, but also poses a series of health risks such as cardiovascular disease—the leading cause of death in 2012, diabetes, musculoskeletal disorders, and "cancer (endometrial, breast, and colon)"[1]. According to the World Health Organization (WHO), in 2014, over 1.9 billion adults (age>18), worldwide, were overweight, which is about 39% of the world's adult population. Among them, 600 million were obese[2].

Historically, most studies were dedicated to the presence and cause of overweight and obesity in developed countries, while few were devoted to this phenomenon in middle- and low-income countries. One reason behind this is that overweight and obesity were more prevalent in developed economies while developing countries were generally suffering from poverty and under-nutrition. Nowadays, developing countries are calling for more attention as many of them are facing a "double burden" of disease: the coexistence of under-nutrition in less developed regions, and a surge of overweight and obesity in more developed regions, and often times this coexistence can be present in the same country and community[3].

The coexistence of under-nutrition in poorer regions and overweight and obesity in richer regions in developing countries presents a weight income relationship that is different than the trend found in developed countries. While obesity and socioeconomic status has an inverse relationship in developed economies, a positive association is usually observed in developing countries: overweight is more prevalent in middle and upper classes who have higher incomes[4]. This phenomenon can be attributed to the two major causes of overweight and obesity and the difference between the levels of economic progress in developed and developing countries.

Overweight is usually caused by an increase in the intake of high-fat, high-caloric food, and a decrease in physical activity due to increased sedentary form of work and transportation[5]. Middle- and high-income population in developing countries have just entered the stage where economic development has brought them enough spare income to shift diet patterns from carbohydrate-concentrated to high-fat, high-protein, energy-dense food. On the other hand, many of them occupy a sedentary-nature job and start to change modes of transportation towards personal vehicles—further decreasing the level of physical activities and increasing the likelihood of being

[1] (Obesity and Overweight: Fact Sheet 2015)
[2] (Obesity and Overweight: Fact Sheet 2015)
[3] (Obesity and Overweight: Fact Sheet 2015)
[4] (Ball and Crawford 2005)
[5] (Obesity and Overweight: Fact Sheet 2015)

overweight. Such lifestyle is usually observed in the low-income groups in developed countries where high-fat and energy-dense food is much cheaper than healthy and nutritional foods. Different from the low-income groups, middle- and high-income population in developed countries have better access to not only healthy diets, but also after-work physical activity facilities. Meanwhile, since most literature was dedicated to studying overweight and obesity in developed countries, population in these areas receive more media coverage on the risks of being overweight and obese, thus are more likely to control fat and sugar intakes and increase physical activity levels to stay healthy. The cultural idea of living a healthy life and the differences in the access to healthy food and exercising facilities explain the negative correlation between income and body weight in developed countries, while little media coverage on such issues in developing countries resulted in a lack of awareness of the risks of overweight and obesity, thus higher income groups, who have better access to variety of food choices, tend to consume more fat and sugar (because they taste good) and are more likely to become overweight.

Even though there have been a plethora of studies and theories in literature about the positive correlation between income and body weight in developing countries, there are few studies on body weight trends over the years, especially regional trends where economic development differs. This paper adds to the literature by not only investigating the general association between income and body weight in developing countries using the case of China, but also focuses on body weight trends over time, as well as the regional and gender disparities in such trends.

II. Background

China's economy went through a major transformation in 1978 during the economic reforms and opening up. The economic reform, accompanied with family planning program and financial accountability within enterprises and service sector organizations brought structural changes to the country's economy. Once one of the poorest countries in the world in 1978, China's economic growth has been unprecedented ever since the reform. Data from the World Bank shows that China's average annual GDP growth from 1978 to 2013 was about 9.85%[6]. According to the IMF, China reached a purchasing-power-adjusted GDP of $17.6 trillion in 2014, surpassed US' $17.4 trillion GDP, and became the world's largest economy.

A rapid economic growth in the past three decades brought productivity increase, higher incomes, and ample food supplies, thus substantially improved Chinese citizens' overall living standards. However, economic improvement is not equally distributed across regions. China's Gini coefficient increased from 27.69 in 1984 to

[6] (World Development Indicators: GDP Growth (annual %) n.d.)

42.06 in 2010, according to the World Bank Gini indicators[7]. The largest share of overall inequality comes from the rural-urban income gaps[8]. Such disparity promoted an expansion of middle and upper classes and substantial living standard improvements in urban areas. However, an increase in living standard also brought more health issues related to increasing body weight due to rapid lifestyle changes in these areas. Fast economic development in urban areas brought Westernization to urban citizens in many aspects, especially diet and transportation. It encouraged a more dramatic diet shift away from the traditional Chinese diet, which composes of high carbohydrate and fiber intakes in the form of rice and cooked vegetables, towards a more westernized diet (such as fast food) with higher fat and caloric intakes. Moreover, higher incomes allow the middle and upper classes to demand more meat and poultry—high-fat and high-protein food that tastes good but was not available on a daily basis before the economic reforms. From the cultural perspective, being overweight is a symbol of wealth in China, thus higher income classes feel highly encouraged to gain weight by demanding more fat and sugar. Moreover, activity levels in the urban setting decrease not only due to a shift from excessive labor work (usually in the countryside) to a more sedentary working nature, but also due to an increased sedentary form of transportation with personal vehicles and convenient public transportation systems.

The data gathered from China Nutrition and Health Survey indicates that from 1991 to 2011, China's overall adult body weight, in terms of Body Mass Index (calculated with weight(kg)/height(cm)^2), experienced a monotone increase in average values (Figure 1). While the median of BMI just increased slightly over the years, heavy upper tail values are observed in the later years, especially 2011, indicating bigger body weight variance and a larger proportion of higher-weight population (Figure 2). In addition, urban areas have a higher average BMI value than rural areas in all those years except 2011 (Figure 3). Figure 4 also shows that urban incomes are generally higher than rural incomes in all the survey years, which is consistent with the claim that high-income groups usually have higher weights than low-income groups.

[7] (World Development Indicators: Gini Indicator n.d.)
[8] (Yang 1999)

SURVEY YEAR	mean(bmi)
1991	**21.67308**
1993	**21.84802**
1997	**22.2986**
2000	**22.86085**
2004	**23.11293**
2006	**23.21118**
2009	**23.4081**
2011	**23.88466**

Figure 1: the overall average BMI value for adults from 1991 to 2011 experienced a monotone increase from 21.67 to 23.88, a 10% increase.

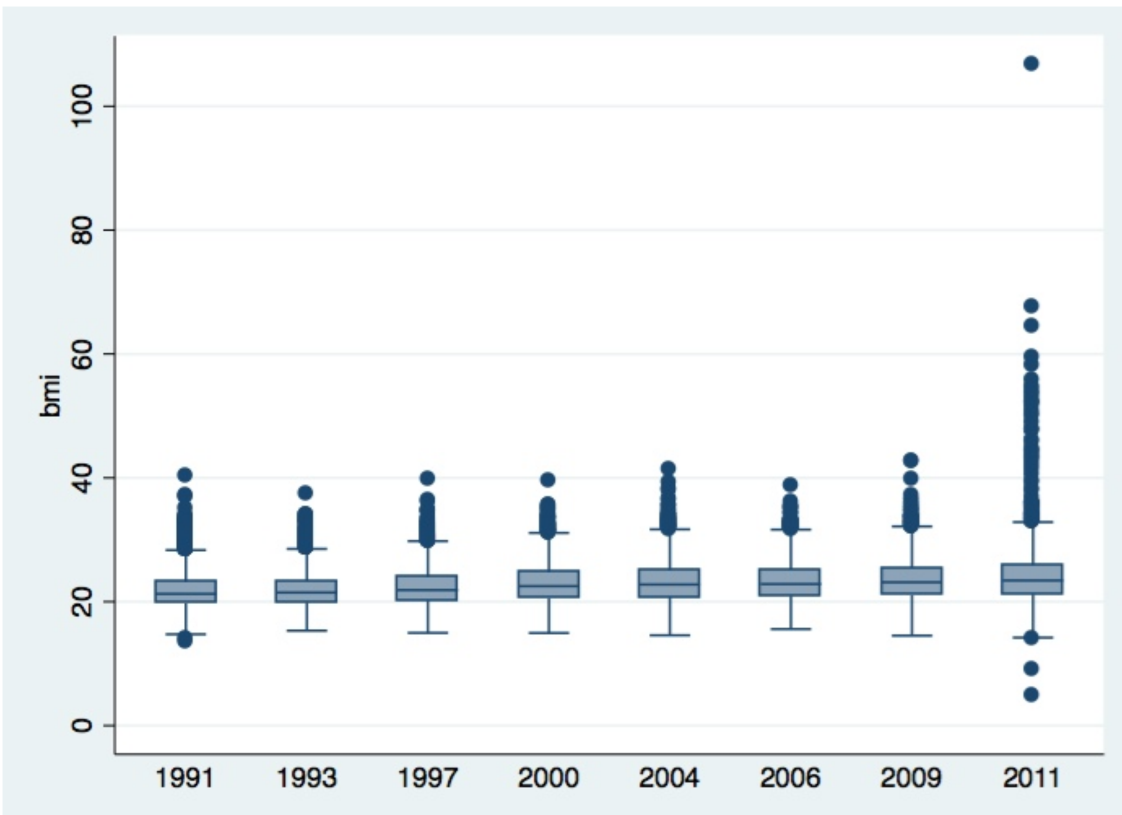

Figure 2: the box plot summarizing the overall adult BMI values from 1991 to 2011. Heavy upper tails are observed in 2011.

SURVEY YEAR	1=URBAN SITE(U) 2=RURAL SITE(R) 1	2
1991	**22.03435**	**21.50643**
1993	**22.21475**	**21.69051**
1997	**22.7123**	**22.09755**
2000	**23.19165**	**22.70783**
2004	**23.34471**	**23.0063**
2006	**23.29597**	**23.17323**
2009	**23.4324**	**23.39764**
2011	**23.7972**	**23.92369**

Figure 3: average adult BMI values by regions. Urban site experience higher BMI averages from 1991 to 2009 while rural site catches up and surpasses urban site in BMI average in 2011.

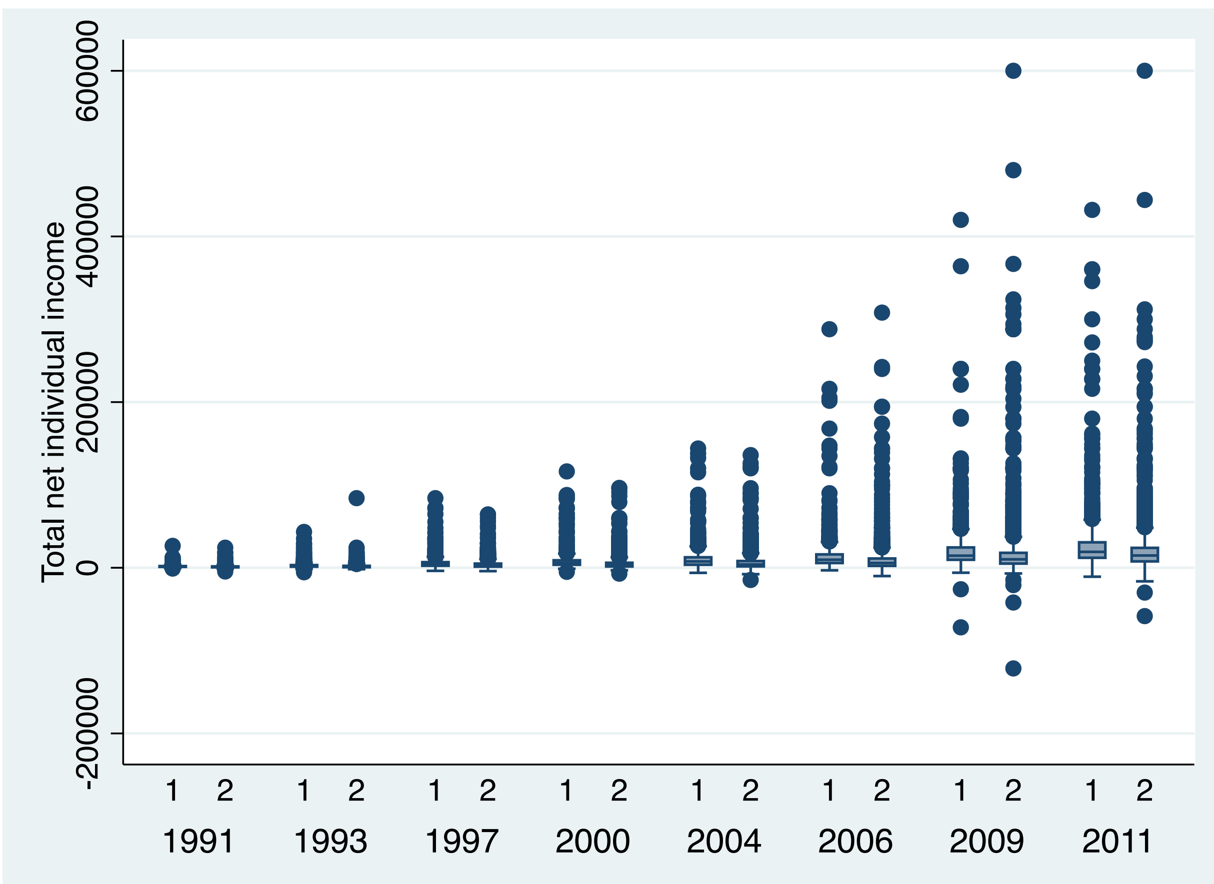

Figure 4: total net individual income from 1991 to 2011, where 1 indicates urban areas, and 2 indicates rural areas.

III. Data

The data set used for this paper is the China Nutrition and Health Survey (CHNS), an international collaborative project between the Carolina Population Center at the University of North Carolina at Chapel Hill and the National Institute of Nutrition and Food Safety at the Chinese Center for Disease Control and Prevention. CHNS data includes nine survey years (1989, 1991, 1993, 1997, 2000, 2004, 2006, 2009, and 2011) with high follow-up levels. The survey used a complex survey design with multistage random cluster sampling to draw about 4,400 households with a total of 26,000 individuals from nine provinces (in dark green and dark blue on Map 1) each year to provide a highly diversified picture of geographic and socioeconomic characteristics. The survey covers costal (Liaoning, Shandong, and Jiangsu, colored in Blue) and mainland (Heilongjiang, Henan, Hubei, Hunan, Guizhou, Guangxi, colored in Green) provinces, as well as rich and poor provinces that vary substantially in economic development, geographic locations, and public resources (see Map 1 for layout).

Map 1: provinces selected for CHNS survey are in dark green and dark blue. A total of nine provinces were selected each year to provide survey samples with diversified backgrounds.

CHNS data set consists of multiple surveys. The master ID survey includes a distinct ID number for each individual surveyed, together with basic information such as gender, age, province, and urban/rural resident indicators. The constructed income survey collects data on annual personal net income from nine different income sources. The physical examination survey collects individual height, weight, and repeated blood pressure measures. In addition to basic physical measures, it also collects information on disease characteristics such as diabetes and high blood pressure. The nutrition survey includes energy intake information such as 3-day average caloric intake, carbohydrate intake, fat intake, and protein intake. The relationship survey contains information on individual's marital status, and the education survey reports individual's educational status from no education to masters' degrees and higher.

The data is longitudinal, so each survey comes with individual IDs and years in which the survey is conducted for the specific person, which makes it easy to map the information from different surveys into one data set according to individual IDs and survey years. Since this paper is interested in studying the relationship between income and adult body weight, only observations with ages from 18 to 60 are included in the final data set. Meanwhile, the nutrition survey was not implemented in the first survey year (1989), in order to be consistent with the rest of the variables of interest, the final data set only contains eight survey years from 1991 to 2011. After proper data cleaning, a total of 47,759 observations were left for further analysis.

IV. Model

The statistical model used in this paper is an OLS regression with multiple dummy variables:

$$BMI= \beta_0 + \beta_1(log(income)) + \beta_2(age) + \beta_3(d3kcal)+ \beta_4(i.education)+ \beta_5(i.marital.status)+ \beta_6(i.coast/noncoast)+ \beta_7(i.urban/rural)+ \beta_8(i.gender)+ \beta_9(i.wave) +\varepsilon$$

where *log(income)* is the log of individual's annual net income, *age* is age recorded in years and *d3kcal* is a 3-day average caloric intake. Variables with "i" in front of them are dummy variables: *i.education* is a dummy variable for education (1=primary school degree and lower, 2=middle school degree, 3=vocational, university, masters, and higher degrees); *i.marital.status* is a dummy variable for marital status (1=single, 2=married); *i.coast/noncoast* is a dummy variable indicating coastal and noncoastal regions (1=coastal regions, 2=noncoastal regions[9]); *i.urban/rural* is a dummy variable indicating urban and rural regions (1=urban, 2=rural); *i.gender* is a dummy variable for gender (1=male, 2=female); and *i.wave* is a dummy variable for all the survey years from 1991 to 2011.

In addition to the OLS regression that incorporates all the observations, separate regressions were run for combinations of gender, costal and noncoastal regions, and urban and rural regions to investigate the gender and regional influences on the regression coefficients.

V. Results and Analysis

The output of the OLS regression based on all observations is summarized in Table 1. The results indicate that BMI is positively associated with income, age, and caloric intakes. The data supports early literature findings that opposite to the negative correlation between income and body weight in developed economies, larger BMIs are more prevalent among higher income groups in developing countries[10]. Moreover, the significant positive coefficient in front of age for Chinese adults is consistent with trends found by Vermeulen, Geomaere, and Kaufman for the Belgium population—BMI is positively correlated with age[11]. Furthermore, the positive association between BMI and caloric intakes supports the claim made by WHO that overweight and obesity are usually caused by an increasing intake of high-caloric food[12].

Gender-wise, the difference between male and female BMI values is not very large, with females having a slightly higher BMI. The data also indicates that middle school graduates have a higher BMI than primary school or lower educational degrees, but the influence of higher education on BMI is unclear. If the same OLS regression is run while separating gender (Table 2 summarizes the regression for males only and Table 3 summarizes the regression for females only), the insignificance from the previous regression can then be explained by the gender difference of the impact of education on BMI: for males, an increasing educational level is associated with higher BMI values, while for females, the opposite is true. The impact of marital status on BMI is also significant: married people tend to gain weight than singled population.

---

[9] Coastal and noncoastal dummy is created based on province geographic locations. Coastal regions include Liaoning, Shandong, and Jiangsu Province and noncoastal regions include Heilongjiang, Henan, Hubei, Hunan, Guizhou, and Guangxi province.

[10] (Ball and Crawford 2005)

[11] (Vermuelen, Goemaere and Kaufman 1999)

[12] (Obesity and Overweight: Fact Sheet 2015)

Linear regression	Number of obs =	42655
	F(16, 42638) =	346.13
	Prob > F =	0.0000
	R-squared =	0.1060
	Root MSE =	3.0787

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.1092394	.0149238	7.32	0.000	.0799885	.1384904
age	.0419297	.0017678	23.72	0.000	.0384648	.0453945
d3kcal	.0000604	.0000212	2.85	0.004	.0000188	.0001019
2.gender	.082261	.0313706	2.62	0.009	.0207741	.1437479
education						
2	.1378228	.0348265	3.96	0.000	.0695621	.2060835
3	.0983259	.0591331	1.66	0.096	-.0175761	.2142279
2.marital_new	.755479	.0508174	14.87	0.000	.655876	.8550821
2.coast_dum	-.9444304	.0324632	-29.09	0.000	-1.008059	-.8808019
2.t2	-.2881775	.0342508	-8.41	0.000	-.3553098	-.2210452
wave						
1993	.1178807	.0475628	2.48	0.013	.0246567	.2111046
1997	.5368914	.0527449	10.18	0.000	.4335103	.6402725
2000	.8866793	.0556908	15.92	0.000	.7775242	.9958343
2004	.9838201	.0608679	16.16	0.000	.8645178	1.103122
2006	.9931254	.0635818	15.62	0.000	.8685038	1.117747
2009	1.095476	.068189	16.07	0.000	.9618238	1.229127
2011	1.4935	.0822624	18.16	0.000	1.332264	1.654736
_cons	19.28813	.1429193	134.96	0.000	19.00801	19.56826

Table 1: OLS regression.

-> gender = 1

Linear regression

Number of obs =	21015
F(15, 20999) =	249.11
Prob > F =	0.0000
R-squared =	0.1384
Root MSE =	3.0018

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.2253835	.0203197	11.09	0.000	.1855553	.2652117
age	.0282656	.0024561	11.51	0.000	.0234516	.0330797
d3kcal	.0000941	.0000314	3.00	0.003	.0000327	.0001556
education						
2	.4951436	.0487864	10.15	0.000	.3995185	.5907687
3	1.0944	.0813987	13.44	0.000	.9348522	1.253947
2.marital_new	.8054582	.0672569	11.98	0.000	.6736295	.9372868
2.coast_dum	-.9298571	.045825	-20.29	0.000	-1.019678	-.8400366
2.t2	-.2601851	.0471842	-5.51	0.000	-.3526698	-.1677005
wave						
1993	.1465866	.062698	2.34	0.019	.0236937	.2694795
1997	.5042636	.0711128	7.09	0.000	.3648771	.6436502
2000	.8745968	.0758655	11.53	0.000	.7258946	1.023299
2004	1.024488	.0830042	12.34	0.000	.8617936	1.187183
2006	1.03747	.0874725	11.86	0.000	.8660174	1.208923
2009	1.140925	.093349	12.22	0.000	.9579537	1.323896
2011	1.576833	.1188302	13.27	0.000	1.343917	1.80975
_cons	18.37699	.1972689	93.16	0.000	17.99033	18.76365

Table 2: OLS regression for males only.

```
-> gender = 2

Linear regression                               Number of obs =    21640
                                                F( 15, 21624) =   181.57
                                                Prob > F      =   0.0000
                                                R-squared     =   0.1050
                                                Root MSE      =   3.1016
```

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.0003616	.0215072	0.02	0.987	-.0417942	.0425174
age	.055553	.0025215	22.03	0.000	.0506107	.0604954
d3kcal	4.68e-06	.0000294	0.16	0.874	-.000053	.0000623
education						
2	-.1025579	.0491029	-2.09	0.037	-.1988033	-.0063125
3	-.8944588	.0823018	-10.87	0.000	-1.055776	-.7331412
2.marital_new	.6328412	.0759087	8.34	0.000	.4840545	.7816278
2.coast_dum	-.9260868	.0453433	-20.42	0.000	-1.014963	-.8372106
2.t2	-.3387624	.0488625	-6.93	0.000	-.4345366	-.2429883
wave						
1993	.0855979	.0707163	1.21	0.226	-.0530112	.2242069
1997	.5535806	.0773172	7.16	0.000	.4020332	.705128
2000	.8652055	.0807912	10.71	0.000	.7068487	1.023562
2004	.8956996	.0877689	10.21	0.000	.723666	1.067733
2006	.888998	.0908444	9.79	0.000	.7109363	1.06706
2009	1.014228	.0981593	10.33	0.000	.8218288	1.206628
2011	1.363821	.1130915	12.06	0.000	1.142153	1.585488
_cons	20.20002	.1940469	104.10	0.000	19.81967	20.58036

Table 3: OLS regression for females only.

Region-wise, noncoastal region residents have a lower BMI than those in coastal regions, and rural residents have a lower BMI than those in urban areas. Coastal and urban areas in China are considered as more developed regions with higher average income and living standards. Higher BMI values in these regions support the claim that a large portion of overweight people reside in more developed regions of the developing countries[13].

Moreover, a monotone increase is observed in BMI values over the years from 1991 to 2011 (all the coefficients in front of the dummy variable *i.wave*, which is summarized in Table 1, is positive and increasing in absolute values). After observing an increasing trend of average and median incomes over time (Figures 5 & 6), such BMI increase can be attributed to an increasing net individual income in the Chinese population.

SURVEY YEAR	mean(indinc)
1991	**1408.316188**
1993	**2018.827318**
1997	**4263.249222**
2000	**5383.525519**
2004	**7088.673928**
2006	**10047.05455**
2009	**16848.7538**
2011	**21745.49355**

Figure 5: Income average by year displays a monotone increase from 1991 to 2011 that is quite large in absolute differences.

[13] (Ball and Crawford 2005)

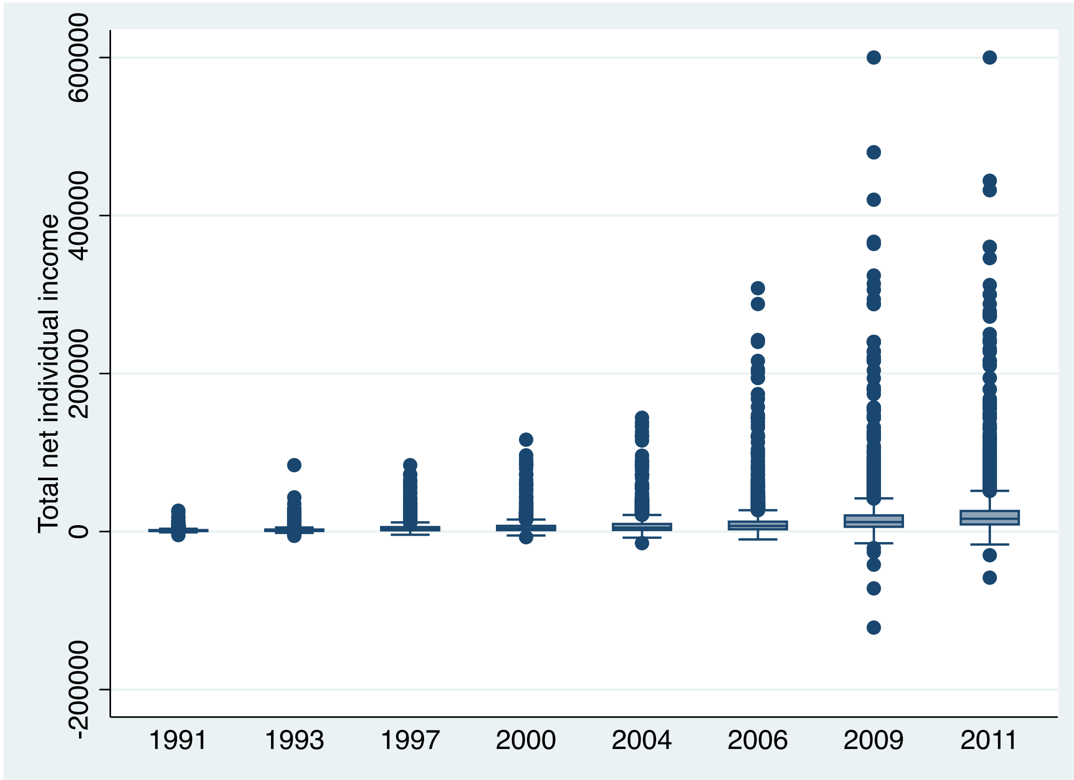

Figure 6: income boxplots over time. Median values steadily increased from 1991 to 2011, while heavy upper tails are more prevalent in later years, indicating existence of more high-income groups.

The previous OLS regression has provided reassuring relationships between BMI and the independent variables of income, age, caloric intakes, gender, region, and time; however, it only shows the gender and regional BMI differences in absolute terms by regression coefficients. In order to investigate the trend in such differences over time, additional regressions separating gender and regions were run, and the results were included in Tables 4-11 in the appendix. The coefficients in front of the year dummies indicate different growth trends for BMIs with the region and gender combinations. Even though the overall trend for all the combinations indicates BMI growth over time, comparisons of the growth trends in each combination tells a story that is more than merely an increasing body weight over the years.

In order to better detect the trend differences for the gender and region combinations over time, all the wave effects are isolated and summarized in Tables 12 & 13. In addition, mean values for each gender and region combination over the years are plotted to better understand how the wave effects and the average BMI values are interacting with one another.

Male, Urban

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.1195829	0.1213403	0.99	0.324
1997	0.506154	0.1349102	3.75	0
2000	0.7659756	0.1432666	5.35	0
2004	0.9078944	0.1565875	5.8	0
2006	0.8206587	0.1682813	4.88	0
2009	0.9238545	0.1826961	5.06	0
2011	1.349058	0.2094138	6.44	0

Male, Rural

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.1592058	0.072657	2.19	0.028
1997	0.4963771	0.0830284	5.98	0
2000	0.9274477	0.0888068	10.44	0
2004	1.074599	0.0975738	11.01	0
2006	1.136152	0.1021439	11.12	0
2009	1.242189	0.1084176	11.46	0
2011	1.704439	0.1434386	11.88	0

Female, Urban

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.1059109	0.1278111	0.83	0.407
1997	0.6902573	0.1370077	5.04	0
2000	0.9048619	0.1431433	6.32	0
2004	0.7395984	0.1571531	4.71	0
2006	0.5631861	0.1650745	3.41	0.001
2009	0.5662235	0.1810658	3.13	0.002
2011	0.8365956	0.2096936	3.99	0

Female, Rural

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.0849834	0.0843946	1.01	0.314
1997	0.4867848	0.0928627	5.24	0
2000	0.8540804	0.0974889	8.76	0
2004	0.9771116	0.1054488	9.27	0
2006	1.04397	0.1084961	9.62	0
2009	1.218852	0.1162269	10.49	0
2011	1.617027	0.1334094	12.12	0

Table 12: wave effects on BMI for gender and rural/urban combinations.

Male, Coastal

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	-0.0818759	0.1113363	-0.74	0.462
1997	0.3861042	0.139343	2.77	0.006
2000	0.7068799	0.1388911	5.09	0
2004	0.8980863	0.148759	6.04	0
2006	0.8654668	0.1597987	5.42	0
2009	0.9918289	0.1709017	5.8	0
2011	1.390871	0.2262362	6.15	0

Male, Noncoastal

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.2707942	0.0750782	3.61	0
1997	0.5531176	0.0822595	6.72	0
2000	0.9521866	0.0902185	10.55	0
2004	1.079736	0.0999456	10.8	0
2006	1.11522	0.104174	10.71	0
2009	1.214554	0.1112217	10.92	0
2011	1.663268	0.1375939	12.09	0

Female, Coastal

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	-0.0140484	0.1219314	-0.12	0.908
1997	0.5656554	0.1487475	3.8	0
2000	0.829788	0.1443549	5.75	0
2004	0.9724736	0.1566877	6.21	0
2006	0.7958385	0.1602303	4.97	0
2009	0.8709389	0.1721463	5.06	0
2011	1.39507	0.2049898	6.81	0

Female, Noncoastal

bmi	Coef.	Robust Std. Err.	t	P>t
wave				
1993	0.1446436	0.0862325	1.68	0.093
1997	0.5462804	0.0906907	6.02	0
2000	0.8782577	0.0974283	9.01	0
2004	0.8341607	0.1056295	7.9	0
2006	0.9234735	0.110559	8.35	0
2009	1.06601	0.1202239	8.87	0
2011	1.30882	0.1353579	9.67	0

Table 13: wave effects on BMI for gender and coastal/noncoastal combinations.

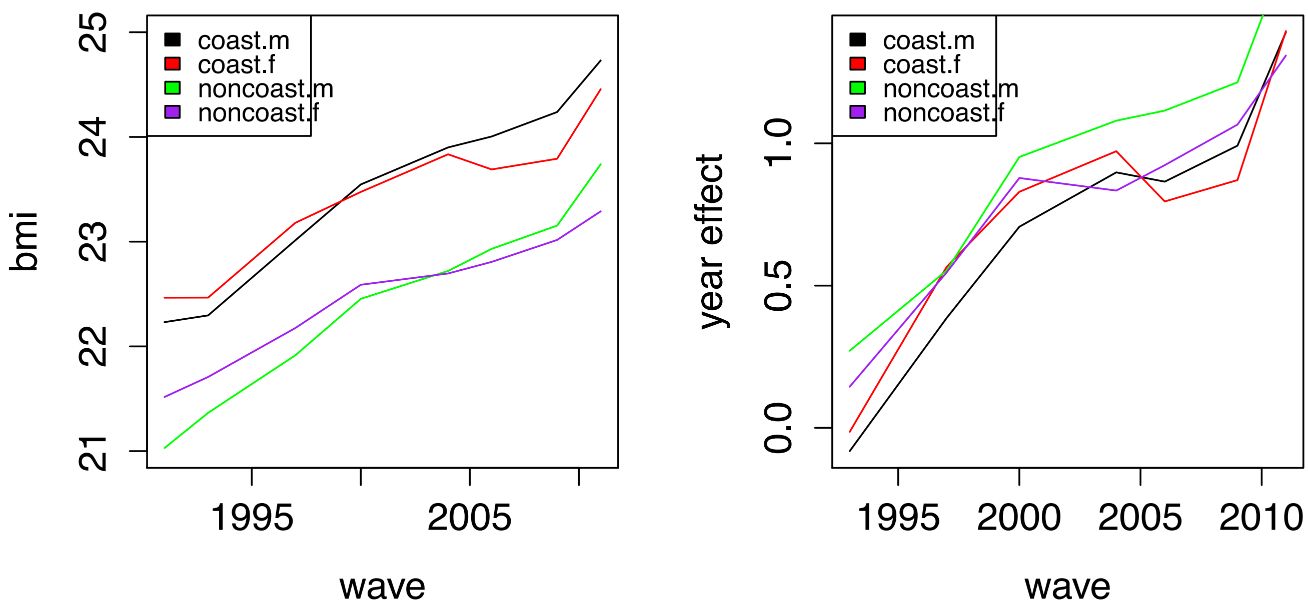

Figure 7: To the left—average BMI values over time for males in coastal areas (black line), females in coastal areas (red line), males in noncoastal areas (green line), and females in noncoastal areas (purple line). To the right—regression wave/year effects on BMI for the same gender and region combinations as the ones to the left.

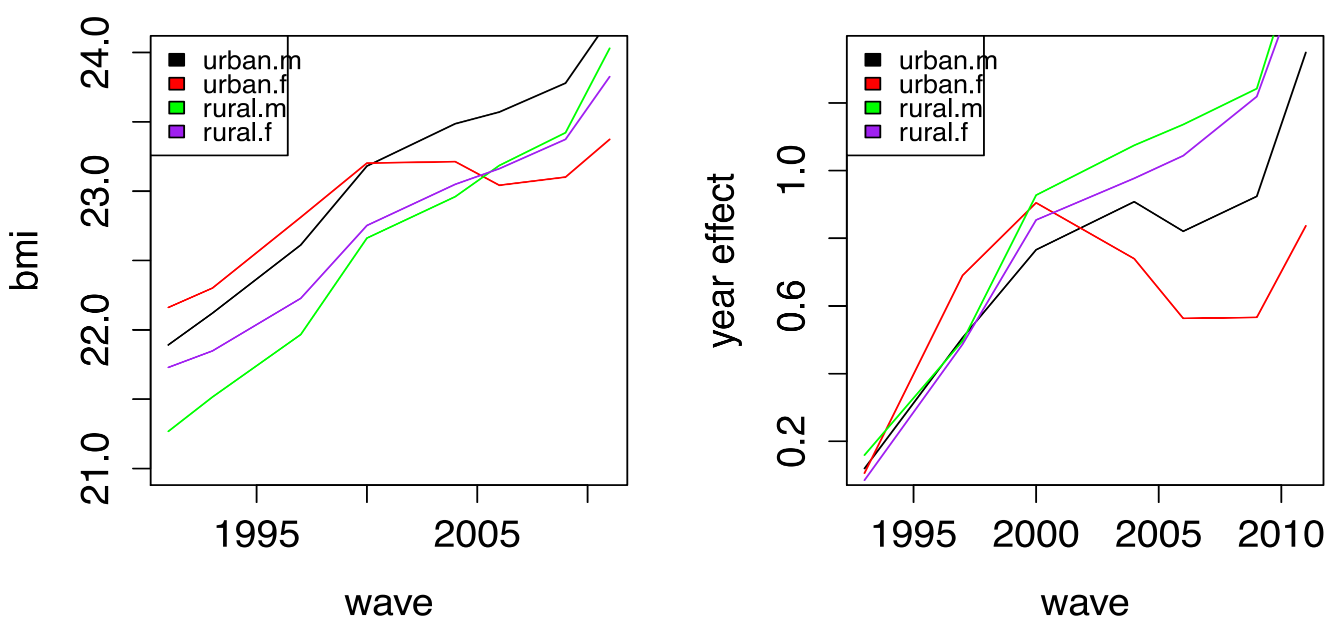

Figure 8: To the left—average BMI values over time for males in urban areas (black line), females in urban areas (red line), males in rural areas (green line), and females in rural areas (purple line). To the right—regression wave/year effects on BMI for the same gender and region combinations as the ones to the left.

Figures 7 and 8 above graphically displayed the average BMI values and the BMI growth (in terms of the year effects) for different gender and region combinations. The average BMI values and regression wave effects indicate an overall increase in Chinese adults body weight over time. More specifically, there are three stages of body weight increase from 1991 to 2011: a rapid increase from 1991 to 2000, and a relatively flatter stage of increase from 2000 to 2009 before it jumps for another steep increase from 2009 to 2011. A possible explanation for the three stages of body weight increase could be the average income increase and economic development over time. The first rapid increasing phase from 1991 to 2011, is merely overcoming hunger. As the economy enjoys the early success of the reforms and opening up policy, average living standard improves and population are able to overcome under-nutrition by consuming more carbohydrate and protein. Such body weight growth trend slows down after 2000 as the under-nutrition problem becomes less prevalent in most areas. As China's economy continues to grow at an unprecedented rate, during the later years, starting from 2009, body weight increase peaks again. The third phase of growth is the affluence phase or the obesity phase. Food consumption shifts from need-based to satisfactory-based, and the trend is driven by an overindulgence of fat and sugar intakes, which results in a larger proportion of overweight and obese people in the overall population (see Figure 9 for graphical reference).

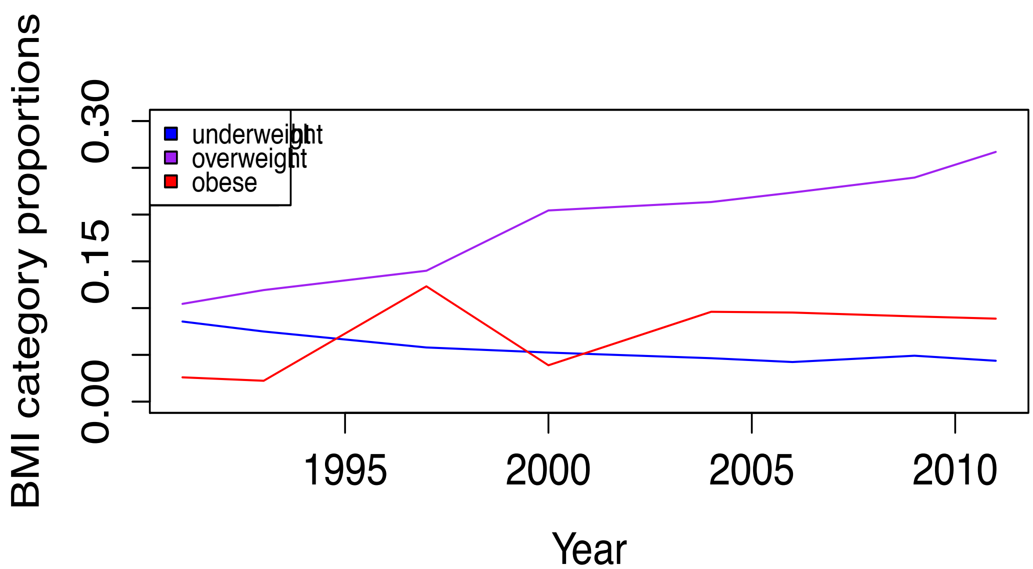

Figure 9: proportion of BMI categories over time. Underweight group refers to a BMI less than 18.5, overweight refers to a BMI between 25 and 30, and obese refers to a BMI over 30.

Due to the increasing proportion of overweight and obese population in the affluence phase, certain disease types that are highly related to overweight and obesity, such as diabetes and high blood pressure, become more common in the general population. Figure 10 displays the adult diabetes ratio in China from 1997 to 2011. The number of diabetes cases in the data set is quite small, which might be because the diabetic indicator in the survey is self-reported, and a lot of times, the survey subject may not be aware of his/her diabetic conditions. However, even though the absolute ratio of the diabetic indicator is small, the trend in this disease over time, especially the steep increase of diabetic proportion since 2004, is consistent with the increasing prevalence of overweight population in the affluence phase.

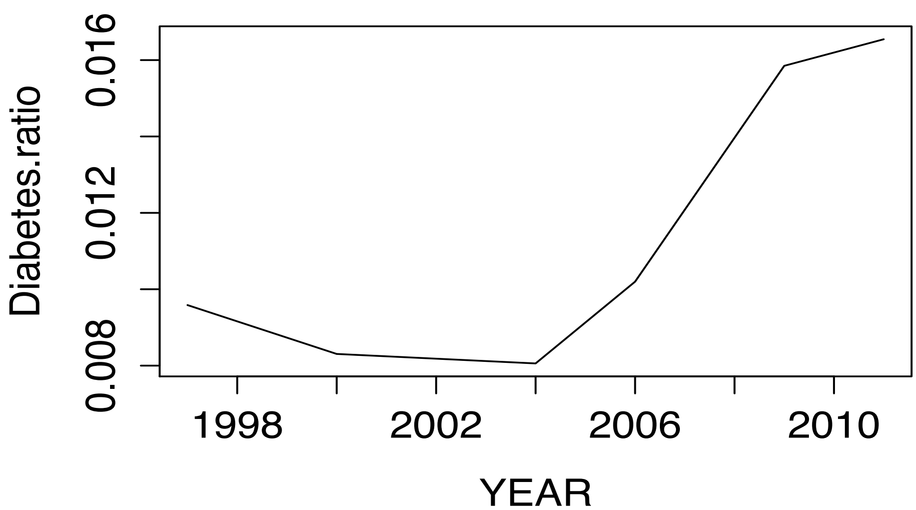

Figure 10: Adult Diabetes ratio, from 1997 to 2011.

Besides diabetes, another risk factor for overweight and obesity is high blood pressure. Figure 11 shows the ratio for the three blood pressure groups (normal, pre-hypertension, and hypertension) over time. From 1991 to 2011, the proportion of population that suffers from pre-hypertension and hypertension steadily increased while the ratio of population with a healthy blood pressure dropped from 60% to almost 30% of the total population during its trough in 2009. The increasing prevalence of health issues that are highly related to overweight and obesity is consistent with the steady body weight increase over time, and indicates the presence of a larger overweight group in the Chinese population.

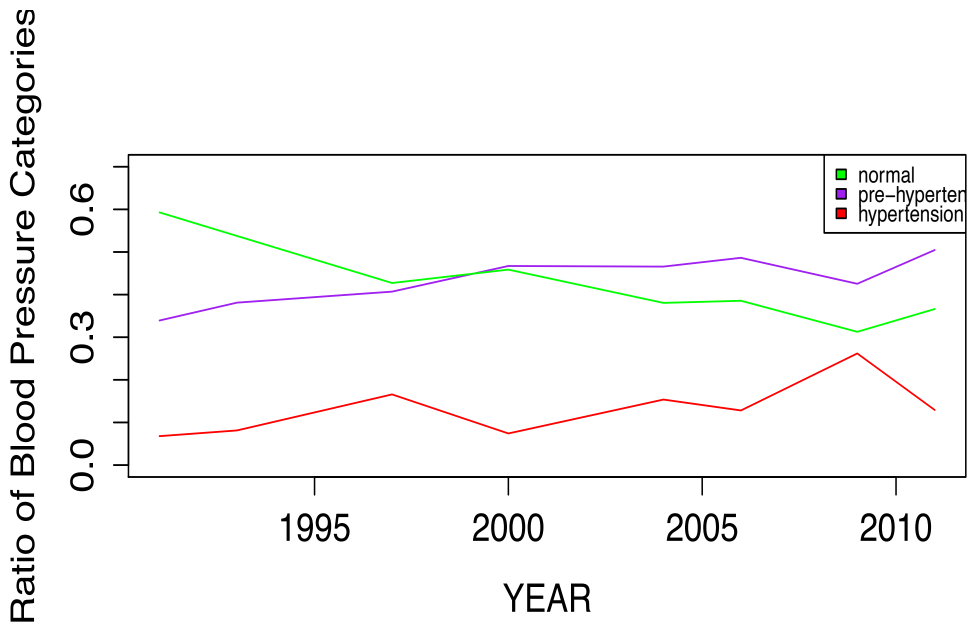

Figure 11: Adult blood pressure category ratios—normal blood pressure refers to a systolic blood pressure of less than 120 mm Hg and a diastolic reading of less than 80; pre-hypertension happens when the systolic reading is between 120 and 139 or the diastolic reading is between 80 and 89; and hypertension is defined when the systolic reading is higher than 140 or the diastolic reading exceeds 90.

Besides the overall body weight increase from 1991 to 2011, regional and gender disparities in the BMI growth trends also exist in the case of China.

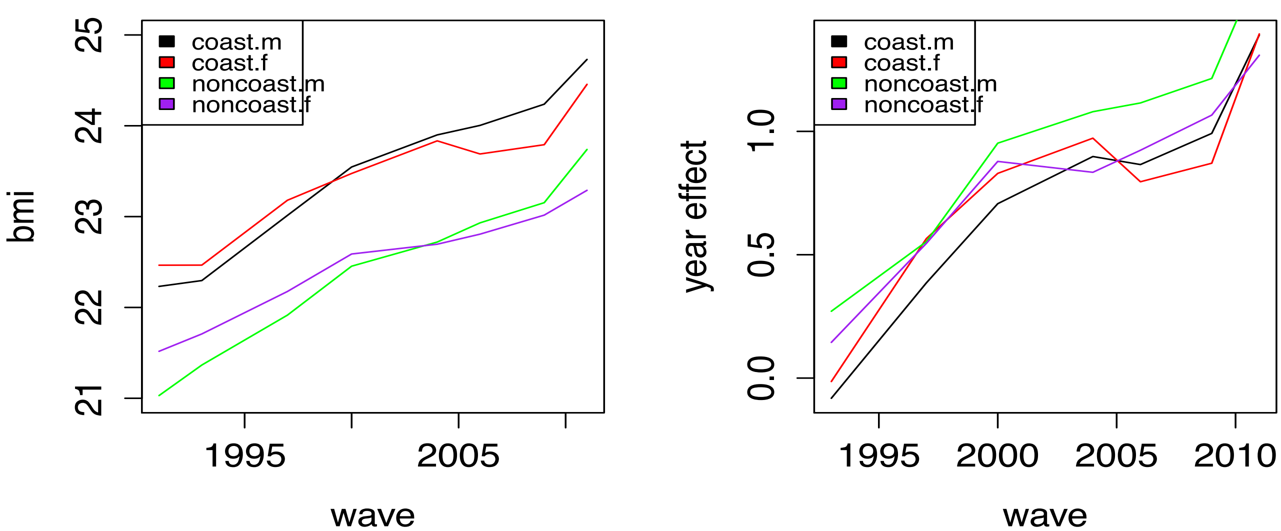

Figure 7: To the left—average BMI values over time for males in coastal areas (black line), females in coastal areas (red line), males in noncoastal areas (green line), and females in noncoastal areas (purple line). To the right—regression wave/year effects on BMI for the same gender and region combinations as the ones to the left.

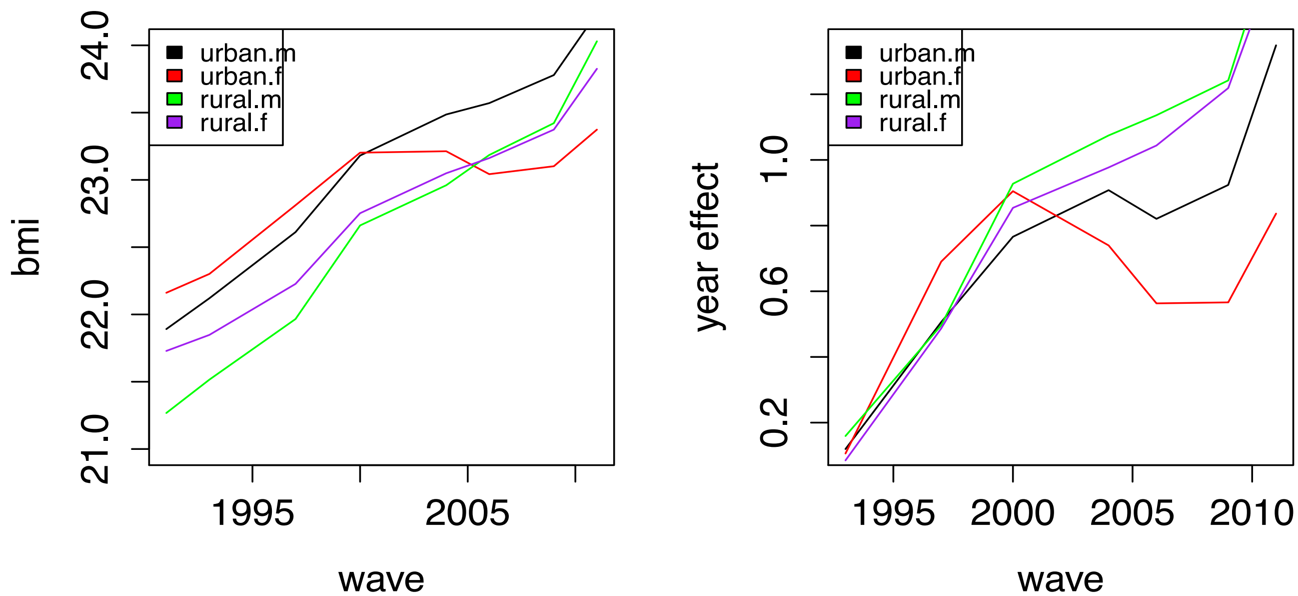

Figure 8: To the left—average BMI values over time for males in urban areas (black line), females in urban areas (red line), males in rural areas (green line), and females in rural areas (purple line). To the right—regression wave/year effects on BMI for the same gender and region combinations as the ones to the left.

When we closely look at Figures 7 and 8 again, we see that average BMI values are higher for coastal regions than noncoastal regions, and also higher in urban areas than rural areas. In China, coastal and urban areas are generally considered to be more economically successful and developed, where average income and living standards are much higher when compared to noncoastal and rural regions. So the higher BMI values in coastal and urban areas are consistent with the theorem that in developing countries, overweight is usually associated with high-income population.

However, the BMI growth trends, in terms of the regression wave coefficients, displays an opposite relationship for coastal/noncoastal areas and urban/rural areas: the growth trend is stronger/higher for noncoastal and rural regions when compared to coastal and urban regions—a body weight catch up experience. The graphical display (Figure 12) of the proportion of overweight and underweight population in urban and rural areas is also consistent with the previous findings: the overweight proportion in urban areas has been higher from 1991 to 2009, and the rural underweight problem has been more prevalent throughout the first half of the survey years. As a result, even though the current average body weight is higher in more developed regions of China, the future growth trend lies in the less developed mainland and rural regions.

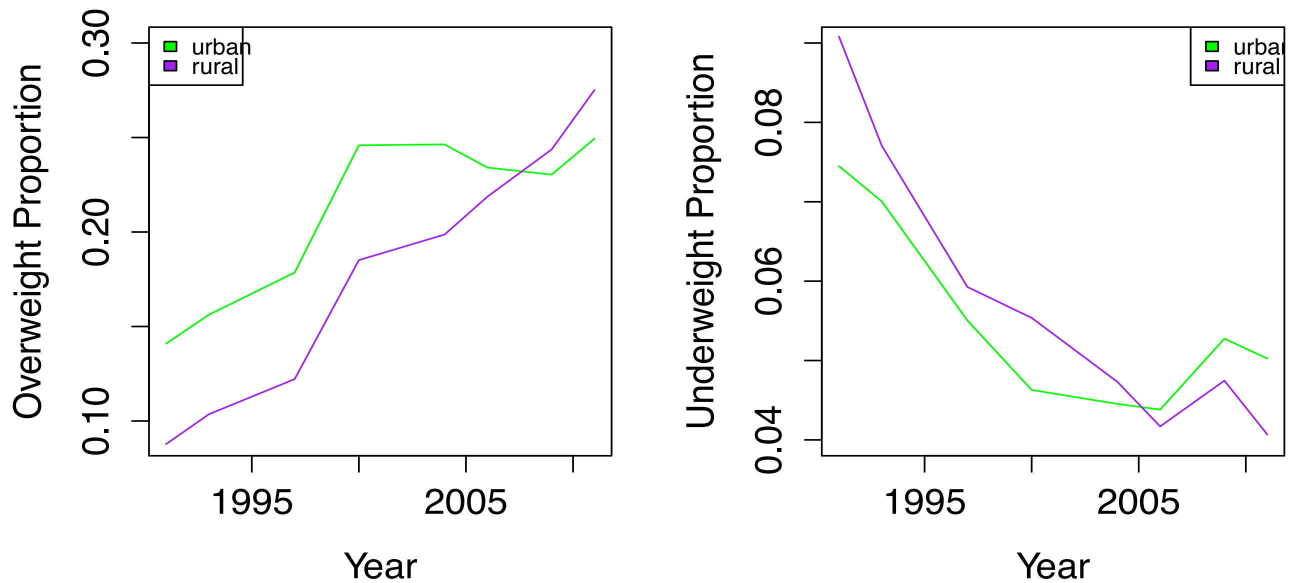

Figure 12: Overweight and underweight proportion in the urban and rural areas.

Gender-wise, females start with higher average BMI values during the first few years of the survey, but end with lower average BMIs towards the later years. This presents another catch-up story between the genders where males have a higher BMI growth trend than females over time. Back to the OLS regression at the beginning of the analysis, the regression results based on all observations (Table 1) indicate a positive relationship between log income and body weight; however, when separate regressions are run for each gender (Tables 2&3), the results indicate that the positive correlation between income and body weight is significant for males only. Given a rapid average income increase over time, it is then reasonable for males to have a much stronger BMI growth trend than females due to the significant correlation between income and body weight. As income rises above a certain level, males are then more likely to become heavier than females.

Other than the overall gender differences in BMI values, the BMI growth trend for urban females stands out as a particular case. Back to Figure 8, urban females displayed a dramatic decrease in BMI growth from 2000 to 2009, and entered a stage of decreasing average BMI values during that time. One possible explanation for the decline in urban females' body weights might be fashion and cultural influences. During the past ten to fifteen years, Chinese culture started to be fond of skinny young women. Media coverage, magazines, the fashion and entertainment industry always feature young women whose BMI values barely reach 18, and thus defined beauty as being underweight. Such a cultural definition of beauty resulted in a large portion of young women constantly putting themselves on diet to lose weight and become “more beautiful”. This phenomenon is more prevalent in urban areas because urban females not only are exposed to more influences from the media due to better access than the rural female population, but also are under deeper modern cultural constraints from their working and rapidly changing living environment.

Statistically, the average BMI values of adult females between the ages of 20 and 35 years old in urban areas experienced a steady increase from 1991 to 1997, a body weight growth possibly due to better nutrition. After 1997, it displayed a dramatic decline in absolute values from its peak at 21.8 in 1997 to almost 21 in 2011 (Figure 13), and further decline after 2011 is highly possible. Figure 14 further indicates that female’s share of the urban underweight population has been steadily increasing ever since 1997, and reached almost 70% of the total underweight population in urban areas in 2011.

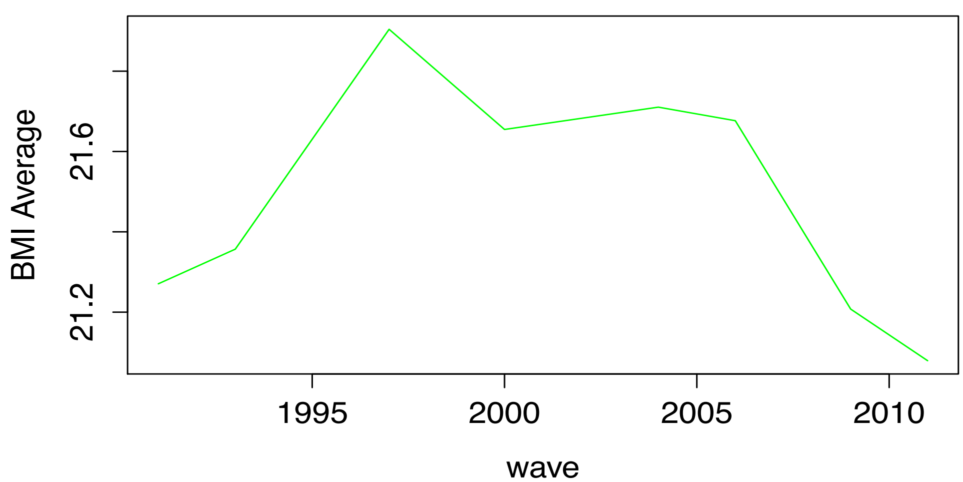

Figure 13: Average BMI values for adult females between the ages of 20 and 35 who reside in urban areas.

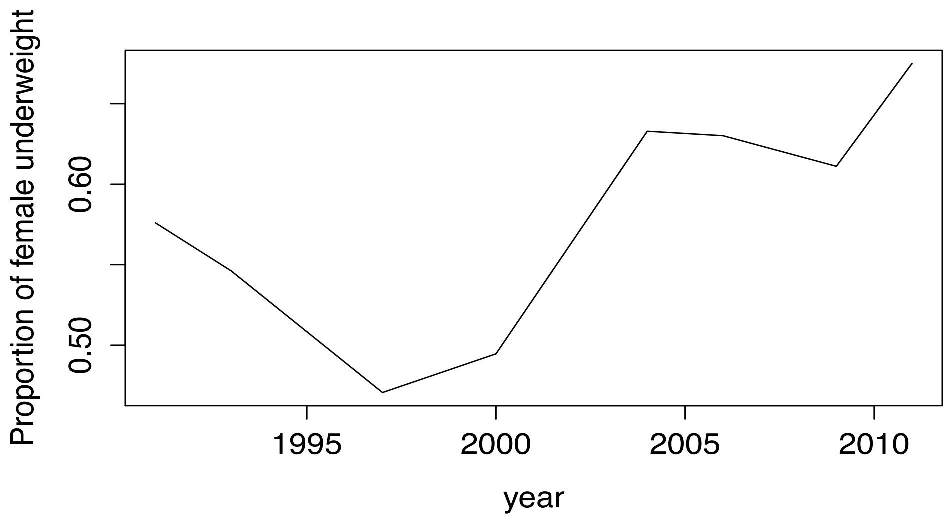

Figure 14: proportion of underweight females among the underweight population in urban areas.

In fact, the problem of being underweight requires as much attention as being overweight and obesity, especially in women. Studies show that underweight and obesity are both associated with excessive death and increased mortality rate[14]. In addition, underweight in pregnant women increases the risk of preterm delivery and low birth weight, which may pose serious risks towards both the mother and the child[15].

The phenomenon of “skinniness=beauty” in urban Chin is actually not unique, it is probably due to the influence of Westernization which brought fashion and cultural trends from the developed countries into China. The Western fashion industry almost exclusively hires underweight models because skinny models are considered more beautiful and elegant. This practice causes many young women who want to join the fashion industry to over-fast on their diet, and many of them develops anorexia nervosa, and sometimes even death due to extreme diets. In early April 2015, The French government passed a new law that bans modeling agents and fashion houses from hiring models who have a BMI value of less than 18, which further states how serious the underweight problem has been in the Western fashion industry. Since the urban areas in China are more exposed to the influence of Westernization, the change in the cultural definition of “beauty” has a much greater impact on its residents than the effect of influence in rural areas, and resulted in a more dramatic decrease in urban females weight trend over the recent years.

---

[14] (Flegal, et al. 2005)
[15] (Siega-Riz, Adair and Hobel 1996)

VI. Conclusions

The empirical analysis of this paper has shown that there exists a positive relationship between income and body weight in China. The BMI trend over time also indicates a three-phase body weight growth along the path of economic development. The first phase of rapid growth is overcoming hunger and under-nutrition due to initial economic success; the second phase is a leveling phase of BMI growth; and the third phase is the affluence phase, where further economic development allows the body weight growth trend to be driven by satisfaction-based fat and sugar overindulgence, and results in overweight and obesity problems.

Regional and gender differences also exist in the developing economy of China. Currently, more developed urban and coastal regions have higher average BMI values than less developed rural and mainland regions, however, with a higher growth trend over time for the less developed regions, future average BMI values are expected to be higher for low-income groups in China. Gender-wise, body weight is significantly related to income for males, thus the body weight growth trend along the economic development path is stronger for males than females, and results in males having higher weights than females as average income increases. Moreover, urban females experience a decline in body weight growth trend for the past 10 to 15 years, probably due to the a change of fashion and cultural identity that "skinniness=beauty", and resulted in an increasing proportion of underweight urban females.

Overall, while the current relationship between income and body weight for Chinese adults is still positive, and overweight and obesity problems are more prevalent in more developed regions, the future body weight growth trend lies in the less developed areas. As the economy keeps growing at a steady and high rate, the relationship between income and body weight is likely to reverse itself in the near future.

Reference

Ball, Kylie, and David Crawford. "Socioeconomic Status and Weight Change in Adults: A Review." *Social Science & Medicine* 60, no. 9 (May 2005): 1987-2010.

Flegal, Katherine, Barry Graubard, David Williamson, and Mitchell Gail. "Excess Deaths Associated With Underweight, Overweight, and Obesity." *The Journal of the American Medical Association*, 2005: 1861-1867.

"Obesity and Overweight: Fact Sheet." *World Health Organization*, January 2015.

Siega-Riz, Anna, Linda Adair, and Calvin Hobel. "Maternal Underweight Status and Inadequate Rate of Weight Gain During the Third Trimester of Pregnancy Increases the Risk of Preterm Delivery." *American Institute of Nutrition*, 1996.

Vermuelen, A, S Goemaere, and JM Kaufman. "Testosterone, Body Composition and Aging." *Journal of Endocrinological Investigation* 22 (1999): 110-116.

"World Development Indicators: GDP Growth (annual %)." *The World Bank.*

"World Development Indicators: Gini Indicator." *The World Bank.*

Yang, Denise. "Urban-biased policies and rising income inequality in China." *American Economic Review*, May 1999: 306-310.

Appendix

```
-> gender = 1, t2 = 1

Linear regression                               Number of obs =      6609
                                                F( 14,   6594) =     75.49
                                                Prob > F      =    0.0000
                                                R-squared     =    0.1212
                                                Root MSE      =     3.093
```

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.232992	.0432607	5.39	0.000	.148187	.317797
age	.0448476	.0045209	9.92	0.000	.0359852	.0537101
d3kcal	.0002322	.0000588	3.95	0.000	.0001169	.0003475
education						
2	.5890839	.1038112	5.67	0.000	.3855803	.7925874
3	.9572931	.1319751	7.25	0.000	.6985791	1.216007
2.marital_new	.6724121	.1260925	5.33	0.000	.4252299	.9195943
2.coast_dum	-.8260174	.0799047	-10.34	0.000	-.9826565	-.6693783
wave						
1993	.1195829	.1213403	0.99	0.324	-.1182834	.3574492
1997	.506154	.1349102	3.75	0.000	.2416863	.7706217
2000	.7659756	.1432666	5.35	0.000	.4851267	1.046824
2004	.9078944	.1565875	5.80	0.000	.6009321	1.214857
2006	.8206587	.1682813	4.88	0.000	.4907728	1.150545
2009	.9238545	.1826961	5.06	0.000	.5657109	1.281998
2011	1.349058	.2094138	6.44	0.000	.9385391	1.759577
_cons	17.41178	.3943302	44.16	0.000	16.63877	18.1848

Table 2: OLS regression for male in urban sites only.

-> gender = 1, t2 = 2

Linear regression

Number of obs =	14406
F(14, 14391) =	183.50
Prob > F =	0.0000
R-squared =	0.1435
Root MSE =	2.9551

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.2174863	.0230914	9.42	0.000	.1722242	.2627484
age	.0218062	.0029137	7.48	0.000	.0160951	.0275174
d3kcal	.0000543	.0000332	1.64	0.102	-.0000107	.0001193
education						
2	.4474424	.0551439	8.11	0.000	.3393534	.5555315
3	1.320311	.1118063	11.81	0.000	1.101156	1.539466
2.marital_new	.8204927	.0791808	10.36	0.000	.6652882	.9756972
2.coast_dum	-.9714587	.055755	-17.42	0.000	-1.080746	-.8621718
wave						
1993	.1592058	.072657	2.19	0.028	.0167887	.3016229
1997	.4963771	.0830284	5.98	0.000	.3336306	.6591235
2000	.9274477	.0888068	10.44	0.000	.7533749	1.101521
2004	1.074599	.0975738	11.01	0.000	.8833418	1.265856
2006	1.136152	.1021439	11.12	0.000	.9359364	1.336367
2009	1.242189	.1084176	11.46	0.000	1.029677	1.454702
2011	1.704439	.1434386	11.88	0.000	1.423281	1.985597
_cons	18.52413	.2123217	87.25	0.000	18.10795	18.9403

Table 3: OLS regression for male in rural sites only.

```
-> gender = 2, t2 = 1

Linear regression                               Number of obs =      6763
                                                F( 14,  6748) =     86.55
                                                Prob > F      =    0.0000
                                                R-squared     =    0.1321
                                                Root MSE      =    3.0623
```

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	-.0111411	.0422674	-0.26	0.792	-.0939986	.0717163
age	.0762609	.0044584	17.10	0.000	.067521	.0850009
d3kcal	.0001574	.0000643	2.45	0.014	.0000315	.0002834
education						
2	-.2349373	.09558	-2.46	0.014	-.4223041	-.0475704
3	-.9965607	.1195124	-8.34	0.000	-1.230843	-.7622786
2.marital_new	.7661883	.1321071	5.80	0.000	.5072167	1.02516
2.coast_dum	-.4841634	.0789514	-6.13	0.000	-.6389331	-.3293937
wave						
1993	.1059109	.1278111	0.83	0.407	-.1446391	.356461
1997	.6902573	.1370077	5.04	0.000	.421679	.9588357
2000	.9048619	.1431433	6.32	0.000	.6242558	1.185468
2004	.7395984	.1571531	4.71	0.000	.4315288	1.047668
2006	.5631861	.1650745	3.41	0.001	.2395879	.8867844
2009	.5662235	.1810658	3.13	0.002	.2112775	.9211696
2011	.8365956	.2096936	3.99	0.000	.4255299	1.247661
_cons	18.93557	.3682378	51.42	0.000	18.21371	19.65743

Table 4: OLS regression for female in urban sites only.

-> gender = 2, t2 = 2

Linear regression

Number of obs	=	14877
F(14, 14862)	=	121.49
Prob > F	=	0.0000
R-squared	=	0.1026
Root MSE	=	3.1025

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.003758	.0249495	0.15	0.880	-.0451462	.0526621
age	.0467473	.0030458	15.35	0.000	.0407772	.0527174
d3kcal	-.0000384	.0000299	-1.28	0.199	-.000097	.0000202
education						
2	-.0838749	.0573465	-1.46	0.144	-.1962811	.0285312
3	-.6547969	.1241705	-5.27	0.000	-.8981864	-.4114074
2.marital_new	.4524665	.0918234	4.93	0.000	.2724813	.6324516
2.coast_dum	-1.113843	.0550286	-20.24	0.000	-1.221706	-1.00598
wave						
1993	.0849834	.0843946	1.01	0.314	-.0804404	.2504073
1997	.4867848	.0928627	5.24	0.000	.3047623	.6688072
2000	.8540804	.0974889	8.76	0.000	.6629902	1.045171
2004	.9771116	.1054488	9.27	0.000	.7704189	1.183804
2006	1.04397	.1084961	9.62	0.000	.8313046	1.256636
2009	1.218852	.1162269	10.49	0.000	.9910328	1.446671
2011	1.617027	.1334094	12.12	0.000	1.355528	1.878526
_cons	20.48675	.213627	95.90	0.000	20.06801	20.90548

Table 5: OLS regression for female in rural sites only.

-> gender = 1, coast_dum = 1

Linear regression

Number of obs =	7111
F(14, 7096) =	73.76
Prob > F =	0.0000
R-squared =	0.1079
Root MSE =	3.1812

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.2315704	.040004	5.79	0.000	.1531507	.3099901
age	.0417053	.0046725	8.93	0.000	.0325457	.0508648
d3kcal	.0001108	.0000484	2.29	0.022	.0000159	.0002057
education						
2	.5597158	.0941778	5.94	0.000	.3750992	.7443325
3	.882559	.1369741	6.44	0.000	.6140489	1.151069
2.marital_new	.8095909	.1288717	6.28	0.000	.556964	1.062218
2.t2	-.2075543	.080741	-2.57	0.010	-.3658307	-.0492778
wave						
1993	-.0818759	.1113363	-0.74	0.462	-.3001282	.1363764
1997	.3861042	.139343	2.77	0.006	.1129502	.6592581
2000	.7068799	.1388911	5.09	0.000	.4346119	.979148
2004	.8980863	.148759	6.04	0.000	.6064744	1.189698
2006	.8654668	.1597987	5.42	0.000	.5522137	1.17872
2009	.9918289	.1709017	5.80	0.000	.6568106	1.326847
2011	1.390871	.2262362	6.15	0.000	.9473808	1.834362
_cons	17.81164	.354367	50.26	0.000	17.11697	18.5063

Table 6: OLS regression for male in coastal areas only.

-> gender = 1, coast_dum = 2

Linear regression

Number of obs =	13904
F(14, 13889) =	149.49
Prob > F =	0.0000
R-squared =	0.1236
Root MSE =	2.903

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.2234348	.023484	9.51	0.000	.177403	.2694666
age	.0213007	.0028381	7.51	0.000	.0157377	.0268637
d3kcal	.0000858	.0000395	2.17	0.030	8.35e-06	.0001633
education						
2	.4603619	.0564911	8.15	0.000	.3496317	.5710922
3	1.228812	.1018188	12.07	0.000	1.029234	1.428391
2.marital_new	.8126936	.078682	10.33	0.000	.6584662	.9669209
2.t2	-.285026	.0580654	-4.91	0.000	-.3988421	-.1712099
wave						
1993	.2707942	.0750782	3.61	0.000	.1236308	.4179577
1997	.5531176	.0822595	6.72	0.000	.3918779	.7143573
2000	.9521866	.0902185	10.55	0.000	.7753461	1.129027
2004	1.079736	.0999456	10.80	0.000	.883829	1.275642
2006	1.11522	.104174	10.71	0.000	.9110254	1.319415
2009	1.214554	.1112217	10.92	0.000	.9965441	1.432563
2011	1.663268	.1375939	12.09	0.000	1.393565	1.932971
_cons	17.71483	.2269429	78.06	0.000	17.27	18.15967

Table 7: OLS regression for male in noncoastal areas only.

-> gender = 2, coast_dum = 1

Linear regression

Number of obs =	7499
F(14, 7484) =	79.73
Prob > F =	0.0000
R-squared =	0.1130
Root MSE =	3.211

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	-.0356876	.039881	-0.89	0.371	-.1138656	.0424905
age	.0677091	.0046659	14.51	0.000	.0585625	.0768556
d3kcal	-4.31e-06	.000055	-0.08	0.938	-.0001122	.0001036
education						
2	-.0636905	.0876369	-0.73	0.467	-.2354834	.1081024
3	-.890802	.1489087	-5.98	0.000	-1.182705	-.5988992
2.marital_new	.9413649	.1271162	7.41	0.000	.6921814	1.190548
2.t2	.0698161	.0872275	0.80	0.424	-.1011742	.2408065
wave						
1993	-.0140484	.1219314	-0.12	0.908	-.2530682	.2249714
1997	.5656554	.1487475	3.80	0.000	.2740684	.8572424
2000	.829788	.1443549	5.75	0.000	.5468119	1.112764
2004	.9724736	.1566877	6.21	0.000	.6653217	1.279626
2006	.7958385	.1602303	4.97	0.000	.4817421	1.109935
2009	.8709389	.1721463	5.06	0.000	.5334837	1.208394
2011	1.39507	.2049898	6.81	0.000	.9932321	1.796907
_cons	19.46221	.3649993	53.32	0.000	18.74671	20.17771

Table 8: OLS regression for female in coastal areas only.

-> gender = 2, coast_dum = 2

Linear regression

Number of obs	=	14141
F(14, 14126)	=	93.21
Prob > F	=	0.0000
R-squared	=	0.0788
Root MSE	=	3.0334

bmi	Coef.	Robust Std. Err.	t	P>|t|	[95% Conf.	Interval]
indinc_log	.0275647	.0257248	1.07	0.284	-.0228594	.0779887
age	.0497967	.0029962	16.62	0.000	.0439238	.0556696
d3kcal	.0000136	.0000353	0.39	0.700	-.0000556	.0000828
education						
2	-.1103695	.0590868	-1.87	0.062	-.2261873	.0054484
3	-.8633909	.0985956	-8.76	0.000	-1.056651	-.6701306
2.marital_new	.445677	.0942423	4.73	0.000	.2609497	.6304043
2.t2	-.5355212	.0592141	-9.04	0.000	-.6515886	-.4194538
wave						
1993	.1446436	.0862325	1.68	0.093	-.0243835	.3136707
1997	.5462804	.0906907	6.02	0.000	.3685147	.7240461
2000	.8782577	.0974283	9.01	0.000	.6872853	1.06923
2004	.8341607	.1056295	7.90	0.000	.6271129	1.041208
2006	.9234735	.110559	8.35	0.000	.7067633	1.140184
2009	1.06601	.1202239	8.87	0.000	.8303557	1.301665
2011	1.30882	.1353579	9.67	0.000	1.043501	1.574139
_cons	19.568	.2237655	87.45	0.000	19.12939	20.00661

Table 9: OLS regression for female in noncoastal areas only.