北京大学中国社会科学调查中心
Institute of Social Science Survey, PEKING UNIVERSITY

CFPS

China Family Panel Studies

中 国 家 庭 追 踪 调 查

Working Paper Series： WP13-003

The Non-Linear Relationship between Body Mass Index and Labor Market Outcomes: New Evidence from China

Mi Luo, Chuanchuan Zhang

2013.10.18

Non-Linear Relationship between Body Mass Index and Labor Market Outcomes: New Evidence from China[1]

Mi Luo[2], Chuanchuan Zhang[3]

First Version: December, 2011
This version: August, 2012

Abstract: Using data from a most recent national household survey in China, we provide new evidence for the associations between body mass index (BMI) and labor market attainments. In contrast to previous studies, we find a significant non-linear relationship between BMI and probability of employment / wages, especially for women. Several potential channels are discussed carefully, including health, occupational sorting, self-esteem, and discrimination. Our findings are in favor of pure discrimination from potential employers, rather than the other three channels.

Key words: Body mass index, Unemployment, Wage, Non-linear correlation

JEL: I12, J31, J64

[1] We would like to thank Yang Yao and the Institute of Social Science Survey at Peking University for sharing the data. All remaining errors are ours.
[2] Corresponding author: Department of Economics, New York University, 19 W. 4th Street, 6FL, New York, NY 10012 (e-mail: mi.luo@nyu.edu, phone: 1-646-201-6109, fax: 1-212-995-4186).
[3] School of Economics, Central University of Finance and Economics (e-mail: zcc@pku.edu.cn). This paper was finished during my visit at Harvard University; I thank Richard Freeman for hosting me.

1. Introduction

What accounts for the labor market attainments has been a long-lasting question. Due to the information asymmetry between the potential employers and employees, the latter try hard to signal their abilities that really matter, for example through education. Since signaling is rarely perfect, discrimination of various kinds might arise.

Previous literature shows that individuals' physical appearances such as beauty, height and weight are related to labor market attainments, though the channels are not yet completely known. For example, Hamermesh and Biddle (1994) found that interviewers' ratings of physical appearances are related to earnings, and the plainness penalty is slightly larger than the beauty premium. Mobius and Rosenblat (2006) tried to use laboratory experiments to figure out the causal effects of appearances on earnings, and they did find a sizable beauty premium.

In terms of body weight in particular, the general finding in the literature is that overweight or obese people, especially women, are less likely to be employed, and are paid lower wages once employed (Harper, 2000; Cawley, 2004; Morris, 2006). Results aiming for causal effects are rather mixed and controversial. For example, a typical study by Cawley (2004) disentangled the correlation between weight and wages. OLS results in that paper indicate negative associations between weight / BMI and wages for white females, black females, Hispanic females and Hispanic males, but positive associations for black males, while IV results suggest significant negative associations only for white females. In Morris (2006), OLS results show a significantly positive relationship between BMI and occupational attainments for males and a significantly negative effect for females. IV results are again not significant. From the literature, we see that at least heavier women tend to have lower attainments in the labor market.

However, these studies generally take as granted the linearity assumption of occupational attainments in weight or BMI. Moreover, earlier studies are usually restricted to either US or European countries, while there is little evidence for developing countries such as China. How is the association like in such an important developing country? And is the linearity assumption necessarily valid? Obviously, if there is indeed a significant non-linear association between BMI and labor market attainments, using linear specifications is likely to give rise to inconsistent estimates. There is even more than that. A non-linear relationship might suggest different mechanisms for the associations on two sides of the optimal BMI level. This kind of asymmetry is worth contemplation itself.

Our study builds upon the literature on the relationship between BMI and labor market attainments for China using the China Family Panel Survey (CFPS), a most recent national household survey, for the first time. We find a significant non-linear relationship, especially for women. In the baseline regression results, the optimal value of BMI for Chinese men in terms of labor market attainments is around 28, and that for Chinese women is around 24.

However, we would like to caution that our results are mere correlations, but not causal relations. Indeed there have been papers that try to dig out a convincing causal relationship, largely with an instrumental-variable (IV) approach. But we shall argue later that the choice of particular instruments, hence the estimation results, are not that valid. The instruments that have been used are essentially which area-based instruments (e.g. Morris, 2006) or a sibling's BMI (e.g. Cawley, 2004), and are likely to fail the exclusion restriction requirement. An exception is probably Behrman and Rosenzweig (2001) who used data for female monozygotic twins and found no statistically significant relationship between BMI and wages. However, Cawley (2004, p. 463) suggested that the insignificance might be due to their small sample size (808).

We try to drive further the associations by exploring potential channels for the non-linear relationship between body mass and labor market attainments. Several potential hypotheses from the literature are entertained. The first possibility is the health channel, i.e. being either underweight or overweight / obese might bring adverse health effects that shall influence performances on the job, thus lowering labor market attainments. Secondly, there might be occupational sorting related to body mass across jobs with different physical requirements, such as between white-collar and blue-collar workers. The third alternative is channeling from BMI to self-esteem, and then to wages. However, we find no evidence supporting the above three channels, hence we think the last channel is more likely, i.e. pure discrimination from potential employers. We also try to provide some evidence for the discrimination channel.

While our robust results call for the importance of considering possible non-linear effects in this literature, non-linearity in the effects of BMI on labor market attainments is in itself intriguing. As the optimal BMI value for both men and women are above the average value in the western countries, we think the non-linear associations exist in the Chinese labor market not merely because it is a developing country and has a less heavy population. Thus non-linearity in the associations should be incorporated in relevant analysis in this literature, and the reasons for the body mass penalty on both sides should be discussed in more detail.

The paper is organized as follows. Following the introduction, the next section discusses data used for analysis and provides some descriptive statistics. Section 3 demonstrates the baseline regression results, and Section 4 discusses in depth evidence for or against potential channels. The last section concludes.

2. Data

The CFPS that we use is by far the largest and latest comprehensive household survey with information on demographic, economic, and health aspects of households in China. It is a biennial survey and is designed to be complementary to the Panel Study of Income Dynamics (PSID) in the United States. The first national wave was conducted under the collaboration of the Institute of Social Science Survey at Peking University and the Survey Research Center at the University of Michigan from April 2010 to August 2010. The five main parts of the questionnaire include communities, households, household members, adults and children data.

The 2010 round covered approximately 14,000 households in 25 provinces, in which 95% of the Chinese population reside. The population is divided into six subpopulation, i.e. five large provinces (Guangdong, Gansu, Liaoning, Henan, Shanghai) and the other 20 provinces. The final sample is made to be representative of 25 provinces through careful weighting.

The survey sample was obtained by three-stage cluster sampling with unequal probabilities. In the first stage, 16 counties were sampled from each of the four large provinces, and 80 counties from 20 other provinces, with probability proportional to population size (PPS). In total there were 144 counties. In the second stage, 2 or 4 administrative villages or resident committees were sampled with PPS in each county or town. Together there were 664 villages or resident committees. In the third stage, 28-42 households were sampled from each village or resident committee, and in all there were about 14,000 households. The national representative final sample covers about 9,500 households and 21,760 adults.

The sample used for analysis in this paper is restricted to those in the labor force, and shall be further restricted to those aged between 18 to 60 years old as working age adults. The two dependent variables capturing labor market attainments are employment status and monthly wage. Covariates to control for in estimating the probability of employment include age, *hukou*[4] status, ethnicity, education attainment, marital status, and self-reported health status. Covariates in the wage equation further include years of working experience and its squared term. There are 8,227 observations in the final sample, though the exact number of observations varies with model specification.

Table 1 presents some summary statistics. In particular, average BMI for Chinese men and women are 23.2 and 21.9, respectively, both lower than American or European counterparts. For example, average BMI of the whole US population is reported as 26.5 in Mocan and Tekin (2009), while that for European men and women are 25.2 and 23.3 in Brunello and Hombres (2007). However, we would like to note that the difference is not that large, and later results suggest that non-linearity comes in not just because Chinese have lower average BMI.

[Table 1 here]

Moreover, 74.2% of men and 72.2% of women in our sample are employed. The numbers are relatively lower probably due to the choice of the range of ages in our sample. The average monthly wage for women is about 600 yuan lower than that for men, and women are also younger on average. Regarding education level, a higher proportion of females graduate from college than males, so women are actually more educated compared to men in our sample.[5] There are about 7.4% of men and 7.9% of women who self-rate a poor health status, which is defined as responding with “not so healthy”, “unhealthy”, or “very unhealthy”. Poor health might influence the earnings,

[4] That is, household registration status.

[5] This may be caused by self-selection of more educated women into the labor force. But China’s 2010 census data does show that women overcount men among recent college graduates. Similar phenomena in OECD countries is sometimes referred as higher education reversal, see (Vincent-Lancrin, 2008).

and we use this dummy as a proxy to control for the potential health channel between BMI and attainment.

Figure 1 provides the lowess graph of probability of employment versus BMI for both genders, and Figure 2 is the corresponding graph for average monthly wages versus BMI. The depictions give us a first sight at how the associations might look like, but we would like to caution that these are without controlling for any potentially influential covariates. That probably partially explains why the associations demonstrated in these two graphs are not perfectly non-linear. We will explore the associations further by regressions in the next section.

[Figure 1 here]

[Figure 2 here]

3. Baseline results

We try to tackle the associations between BMI and labor market attainments by OLS in the baseline regressions. A prototypical regression is:

$$y_{ij}=\beta_0+\beta_1 BMI_{ij}+\beta_2 BMI_{ij}^2+\beta_3 X_{ij}+\alpha_j+\epsilon_{ij} \quad (1)$$

where βs are parameters to be estimated and y_{ij} is a measure of labor attainment (a dummy indicating being employed or not, or a continuous value of monthly wages) for an individual i in county j. BMI_{ij} is the BMI value, defined as the weight (in kilograms) over height (in meters) squared.[6] In addition, X_{ij} is a vector of individual-level control variables for personal characteristics and parents' information. α_j is county fixed effects, and ϵ_{ij} denotes the error term, which is assumed to be mean independent of the BMI value and the control variables for OLS regressions. We are mostly interested in the coefficients before BMI and its squared term, though the other coefficients tell a lot as well.

We first present estimation results for employment probability in Table 2. Results in Columns 1, 3, 4 consistently suggest that men's probability of employment significantly increases with BMI for lower values, but decreases after a cutoff point.

[6] We also categorize weight status according to the WHO classification, i.e. underweight for BMI below 18.50, normal range for between 18.50 and 24.99, overweight (but pre-obese) for between 25.00 and 29.99, obese for over 30.00.

Column 1 only controls for personal characteristics as listed in the table, while Column 3 adds parental characteristics such as parents' education levels and whether any parent is a manager. Column 4 further controls for county fixed effects. Controlling for parents' characteristics and county fixed effects only increases the significance of the effect. For instance, the baseline result in Column 4 indicates that the optimal level of BMI for men in terms of employment probability is around 27.71.[7] And deviating from this value to each side will decrease the probability of getting a job.

[Table 2 here]

Moreover, Column 2 aims to figure out which side of the optimal BMI level is driving the non-linear associations. While the coefficient before BMI decreases, none of the three category dummies is statistically significant. This is consistent with the squared term of BMI not significant for the male sample. Apparently the curvature is not very large.

The last four columns display parallel robust and more significant results for women. The larger coefficients before BMI in Columns 5, 7, 8 for women suggest that women's employment is more sensitive to BMI than for men. From the baseline results in Column 8, the optimal value of BMI for women is around 23.84. And results in Column 6 indicate that being obese or overweight will significantly lower women's probability of getting a job, while being underweight does not carry a significant influence.

On the other hand, the coefficients before age and its squared term also indicate that age has a quadratic association with probability of being employed, though the relationship is statistically significant for women but not for men in almost all the specifications. Having an urban *hukou* boosts the employment prospects for men significantly but not for women, while being a Han is good for both genders. A positive gradient for education level and probability of employment is not surprisingly present for both genders. Being married, compared to not married, is also good for men but has no significant influence on women, while being separated or widowed is no different from the base group of being not married. Noticeably, a self-rated poor

[7] The optimal value of BMI is calculated as the coefficient before BMI divided by (2* absolute value of the coefficient before BMI squared). We use more accurate numbers for calculating the optimal value, so directly reading from the table here might lead to different results.

health is bad for both genders significantly, and the influence is even larger for women. Among parents' characteristics, only mother's education matters, and the influence is weaker for women compared to men.

In a nutshell, results for men and women both suggest a non-linear effect, in contrast with all previous studies which either found no significant effects or linear negative effects. To the best of our knowledge, few previous studies take into account of the possibility of a non-linear relationship, with some exceptions such as Mocan and Tekin (2009) considering quadratic terms of BMI but finding insignificant effects.

We calculate the effects of BMI on employment for various values of BMI based on coefficients displayed in Column 4 and Column 8 of Table 2, for both genders respectively. The pattern is depicted in Figure 3. Turning points for both genders are higher than average values in western countries, which suggests the increasing part of the non-linear effects does not come solely from a lower average BMI in developing countries.

[Figure 3 here]

We further examine the effect of BMI on monthly wages of currently employed workers. Table 3 shows the effects of men's BMI on log value of monthly wages are all insignificant across various model specifications. In particular, Column 1 includes only individual characteristics as covariates. Column 3 further controls for parents' information as in Table 2 and also county fixed effects, and Column 4 controls for occupation type dummies in addition. Column 2, in parallel to Table 2, leaves out BMI squared but includes three weight status dummies, i.e. underweight, overweight but pre-obese, and obese.

[Table 3 here]

In contrast to men, effects for women are highly significant and non-linear. Based on the results in Column 8, the optimal level of BMI for women in terms of wages is around 24.19, which is close to the turning point in estimating the employment probability. We calculate the effects of women's BMI on wages based on coefficients in Column 8 of Table 3 and plot them in Figure 3. However, results from Column 6 shows that it is the underweight penalty that is driving the non-linear associations for women's BMI and wages, which is different from Column 6 in Table 2. This contrast partially implies that the underlying mechanism for the non-linearity for women could be different for getting a job and obtaining higher wages once employed. We will come back to this point in further detail in the next section.

Moreover, age seems to be irrelevant for monthly earnings for men and women, once experience and its squared term are controlled for. Experience has an inverse U-shaped association with earnings, too. Having an urban *hukou* and being a Han do not have influence on wages for both genders. Higher education, compared to being illiterate, always increases wages for men, but not for women. Surprisingly, only obtaining a college degree or above boosts up women's wages. Being married is good for men but bad for women, in terms of wages, while being separated or widowed is even worse for women compared to being not married. A poor health has a negative impact for both genders' earnings, though the effect is not statistically significant for women. Furthermore, being in different occupation types matter a lot for men's wages, but not significantly for women.

The above are baseline results from OLS regressions. We would like to pause for a second here and discuss why we did not go with an IV approach. As mentioned earlier, the previous literature usually chooses BMI of a sibling or area-based average BMI value for instrument. We think both alternatives are likely to reach an inconsistent estimate. On the one hand, a sibling's BMI is closely related to an individual's family background, thus fails the exclusion restriction. For instance, if an individual is born from a richer family, her sibling's BMI might be higher and her labor market attainments are likely to be higher as well. The IV is therefore not completely uncorrelated with the error term. On the other hand, an area-based instrument could suffer from similar problems. A higher average BMI value in the local community might suggest better development of the region, thus a person from this area might be more likely to have better occupational prospects. Since area-based instruments are from higher level as compared to sibling's information, they might suffer from less endogeneity issue.

We have actually tried an area-based IV with our sample, the average BMI value for the county where an individual dwells. As expected, the coefficients become not so credible – they become very large[8], although the inverse U-shaped associations still remain and the statistical significance holds as well. It is also worthwhile to mention

---

[8] Results are omitted here due to space limit, but are available upon request. For instance, in the wage equation for women, a typical estimate for BMI is around 5.067, while that for BMI squared is -0.107. Noticeably, the cutoff point is now 23.68, which is close to the one in OLS regression.

that coefficients for a quadratic term need to be read together, hence larger coefficients do not necessarily mean larger effects. That said, we are still satisfied with OLS results for the moment, and take IV results only as a robustness check.[9]

4. Mechanism

The previous section hopefully establishes sufficient evidence that there is a non-linear association between BMI and labor market attainments in China's labor market, especially for women. In this section, we aim to dig further some potential channels in the literature that could explain the associations. We would like to emphasize again that we are looking at correlations rather than causal effects.

4.1 Health channel

The first apparent channel is through health. Being underweight suggests malnutrition, while being overweight or obese might lead to a series of chronic diseases such as hypertension, heart diseases, or even stroke. Extremely low or extremely high BMI value suggests poorer health as compared to the normal range of BMI, and poor health might adversely influence the productivity at work, thus individual labor market attainments. We first try to omit the poor health dummy from the baseline regression, and repeat the otherwise same specifications. The coefficients before BMI and its squared term are very close to the baseline results. For instance, in the wage equation for women following the specification in Column 7 in Table 3, the coefficients now turn out to be 0.419 and -0.009, respectively.

We then split the whole sample into people with a self-rated good health (reporting "very healthy" or "healthy") versus people with a self-rated poor health (reporting "not so healthy", "unhealthy", or "very unhealthy"), and repeat the baseline regressions on the subsamples in turn. The results are reported in Table 4.

[Table 4 here]

[9] A more ideal instrument would be some exogenous policy changes, historical events or geographical differences that influence individuals' weight or height, thus BMI, but do not affect labor market attainments. Since our sample is cross-sectional, it is relatively more difficult to find such an instrument. We might want to wait till more data to build up a panel structure, in order to try some more valid instruments.

Columns 1-4 of Table 4 are for the results with employment status as the dependent variable, and Columns 5-7 are for the wage equations. The odd-numbered columns are for the healthier subsample, while the even-numbered columns are for the remaining less healthy subsample. Since there are much fewer observations reporting poor health in comparison to good health, and also because we are interested in the association between BMI and labor market attainments through the health channel, we would like to focus on results from the healthier subsample. We see that the inverse U-shaped non-linear associations largely remain, but the results are only statistically significant for healthier women. Again, the insignificance of coefficients for people with poorer health might be partially due to the much smaller sample size.

The result that a non-linear association remains between BMI and labor market attainments even for healthy people suggests that health is not the main underlying mechanism.

4.2 Occupational sorting

Previous studies hypothesized that effects of BMI on wages may come from occupation sorting, or have roots in the different physical requirements for production in different occupations (Harper, 2000). Some studies do find heterogeneous effects of physical appearances on labor market attainments (Hamermesh and Biddle 1994; Harper, 2000). In our case, results in Column 8 of Table 3 with occupation dummies controlled for are almost the same as those in Column 7, which suggests that occupational sorting might not be that important.

Moreover, we estimate the wage equation for white-collar workers and blue-collar workers[10] of both genders in Table 5. Estimates for men are still insignificant, while estimates for women are significant for both white-collar and blue-collar workers. Different magnitudes in Columns 2 and 4 suggest subtle differences across occupations, though a formal *t* test does not reject the equality between two coefficients statistically. The slightly larger effect in the blue-collar group may suggest that BMI is related to productivities somehow, particularly for

[10] White-collar workers are defined as government / firm / enterprise officials, high-skilled workers, or administrative staff. Blue-collar workers are defined as those in the service, agriculture, manufacture industries. There are three other categories that we left out in this exercise: laid-offs, other occupations, or unknown occupations. We also exclude those in the army in our sample.

women. But the effect of BMI is still large and significant effect of BMI in the white-collar group. Therefore, there is no strong evidence that the non-linear associations between BMI and labor market attainments are due to occupational sorting in jobs with different physical requirements.

[Table 5 here]

4.3 Self-esteem

Another alternative explanation suggested by Mocan and Tekin (2009) is that wages are influenced by obesity through the channel of obesity to self-esteem, then to wages. We test this argument in Table 6 but find no significant correlations between BMI and self-esteem, although self-esteem is positively related to wages.[11]

Three variables are chosen as a proxy for self-esteem, which are all based on self-rated scores for certain perspectives of self-esteem. We define the dependent variables as a dummy that equals one if an individual's self-rating exceeds three out of five, and zero if otherwise. Thus we are assuming that the higher the score is, the higher an individual's self-esteem is. The first question, corresponding to Columns 1 and 2 for men and women, is "How popular do you think you are among others?" The second question, corresponding to Columns 3 and 4, is "How well do you get along with others?" The third question, corresponding to the last two columns, is "How confident are you about your future?"

[Table 6 here]

Throughout the six columns in Table 6, both BMI and its squared term are statistically insignificant, and the economic significance is not big, either. Relatively speaking, the impact of BMI is bigger on self-esteem of women than men, which is consistent with previous comparison. From this exercise, it is hard to conclude that self-esteem is an important channel for explaining the strong non-linear associations between BMI and labor market attainments.

4.4 Pure discrimination

[11] The results for the positive relationship between self-esteem and wages are omitted here due to space limit, but available upon request.

We come to the last but probably most important channel in this subsection, i.e. pure discrimination from potential employers. Employers might *a priori* have discrimination against people with either very high or very low BMI value, and have preference for those in the normal range for various reasons. For instance, the employers might think people with a normal-range BMI are healthier, thus can achieve better performance on the job. Please note that this channel is different from the previous health channel, since through the health channel it means that people with a normal-range BMI are indeed healthier and can have higher on-the-job performance, but through discrimination channel the non-linear associations exist because the employers have such *a priori* beliefs. There might be other reasons for discrimination, such as a discrimination effect for better physical appearances according to Hamermesh and Biddle (1994).

Actually, some of the results from previous discussions on the other three potential channels are already suggesting the possibility of pure discrimination. The fact that the non-linear associations remain for the healthier subsample and the white-collar workers indicates that health and occupational sorting due to physical requirements cannot explain the associations, and the self-esteem is also not related to BMI. The only remaining possibility is discrimination. Moreover, BMI and its squared term are much more significant in determining women's wages but not men's，might also suggest that discrimination on the job is more serious for females.

Interestingly, from Column 6 in Table 2 and the same column in Table 3, we see that the discrimination effect is stronger on the overweight / obese side rather than underweight side for women before being employed, yet the pattern switches to the opposite once the individual is employed. There are several possible explanations for this switch. One possibility is that employers asymmetric preferences before and after recruiting new hands. Once very heavy individuals are already excluded from the recruiting pool and the average BMI of employees is lowered, then employers have preferences for people with a larger BMI given this new pool.

5. Conclusion

This study re-examines the relationship between BMI and labor market outcomes using data from a most recent nationally representative household survey in China.

The results reinforce the existing literature that BMI has significant impacts on employment status and wages, especially for women. However, in contrast to almost all previous empirical studies, effects of BMI on both employment and wage are non-linear, suggesting being either overweight or too skinny brings a penalty on labor market outcomes. Our results are very robust to different specifications, thus call for the importance of considering non-linear effects in this literature, as leaving out the non-linear term might lead to inconsistent estimates.

We also look at potential channels of this non-linear effect, and find some suggestive evidence for pure discrimination and against health, occupational sorting, and self-esteem, though more substantial evidence is needed in future to be conclusive. The contrast of our results for China and those in previous literature on western countries – if non-linear effects indeed do not exist for the latter – might come from the gap in average BMI between developing countries and developed ones, though the fact that turning points of non-linear effects are higher than average BMI in western countries seems to provide some counter evidence for this claim. Fundamental differences between the labor markets of developing and developed countries, such as in what way discrimination takes place, in this aspect of BMI are worth further investigation.

References

Behrman, J., and M. Rosenzweig. 2001. The returns to increasing body weight. PIER Working Paper No. 01-052.

Brunello, G., and B. Hombres. 2007. Does body weight affect wages: Evidence from Europe. *Economics And Human Biology* 5: 1-19.

Cawley, J. 2004. The impact of obesity on wages. *Journal of Human Resources* 39 (2): 451-474.

Hamermesh, D., and J. Biddle. 1994. Beauty and the labor market. *American Economic Review* 84 (5): 1174-1194.

Harper, B. 2000. Beauty, stature and the labor market: a British cohort study. *Oxford Bulletin of Economics and Statistics* 62: 771-800.

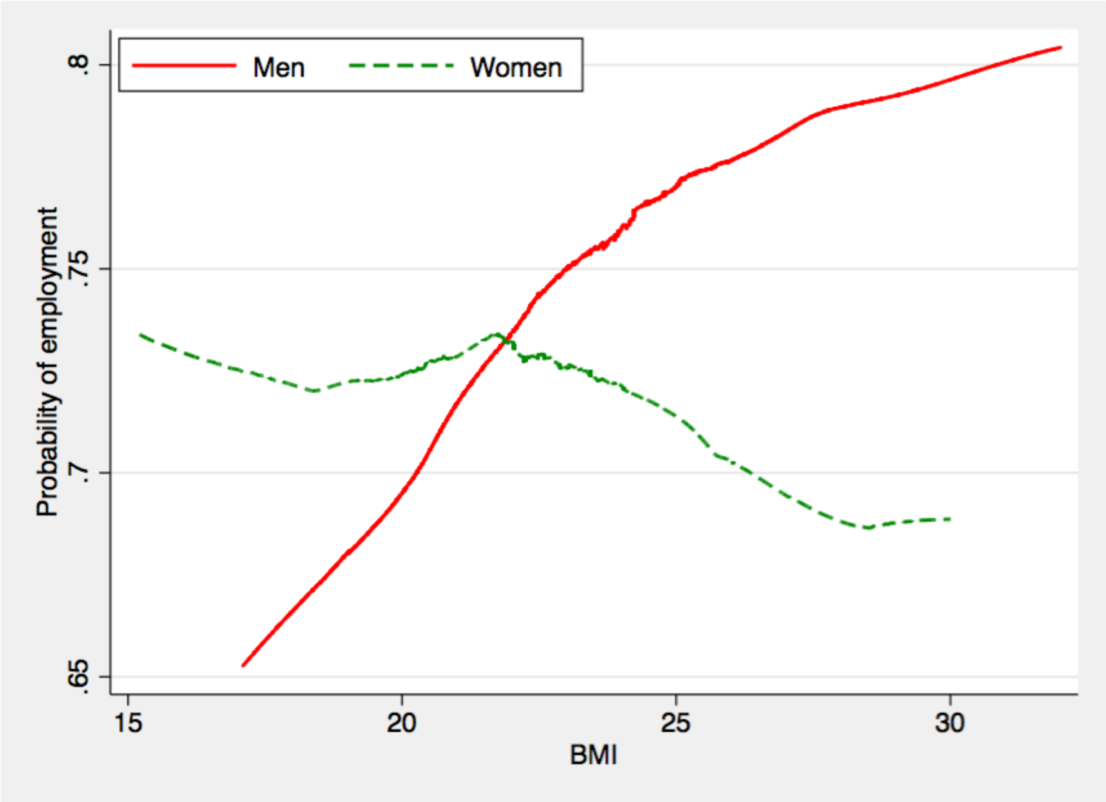
Mobius, M., and T. Rosenblat. 2006. Why beauty matters. *American Economic Review* 96 (1): 222-235.

Mocan N., and E. Tekin. 2009. Obesity, self-esteem and wages. NBER working paper, No. 15101.

Morris, S. 2006. Body mass index and occupational attainment. *Journal of Health Economics* 25: 347-364.

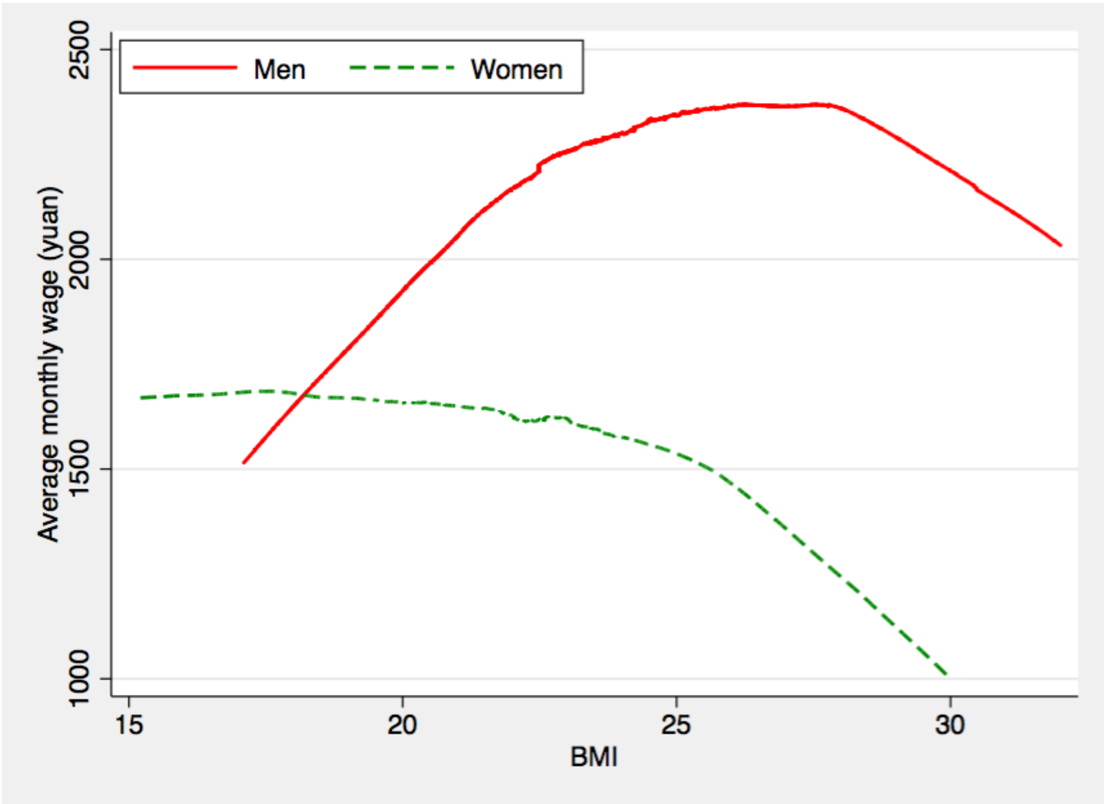
Vincent-Lancrin, S. 2008. The reversal of gender inequality in higher education: an on-going trend. *Higher education to 2030: demography,* Vol. 1, Chapter 10, OECD.

FIGURE 1 Lowess of probability of employment vs. BMI



Notes: The sample for drawing the lowess graph is restricted for BMI between 17 and 32 for men, and between 15 and 30 for women. The two cutoff points on each side are chosen by cutting an approximately 1% tail on both sides. We would like to caution that on each side of extremely small or extremely large BMI value, the estimation would not be very accurate as there are fewer observations compared to in the middle. The bandwidth is by default 0.8.

FIGURE 2 Lowess of average monthly wage vs. BMI



Notes: The sample for drawing the lowess graph is restricted for BMI between 17 and 32 for men, and between 15 and 30 for women. The two cutoff points on each side are chosen by cutting an approximately 1% tail on both sides. We would like to caution that on each side of extremely small or extremely large BMI value, the estimation would not be very accurate as there are fewer observations compared to in the middle. The bandwidth is by default 0.8.

FIGURE 3 Effects of BMI on probability of employment and average monthly wages

Notes: This figure demonstrates the marginal effects of BMI and its squared term on the probability of being employed for men and women, and log value of monthly wages once employed for women only. The solid green line is for employment probability of men, the dashed red line is for employment probability of women, and the dotted black line is for the wage equation of women. The left vertical axis is for the employment probability regression, while the right vertical axis is for the wage equation. The vertical lines denote the turning points from a positive effect of increasing BMI to a negative one, corresponding to three regression specifications respectively.

TABLE 1 Descriptive statistics

	Men		Women	
Variable	Mean	Std. Dev.	Mean	Std. Dev.
BMI	23.22	3.278	21.85	2.975
% employed	0.742	0.437	0.722	0.448
Monthly wage	2217	2371.5	1616.5	1578.4
Age	39.21	11.25	36.45	10.19
Education levels (%)				
Illiterate	0.073	0.261	0.088	0.284
Primary school	0.148	0.356	0.146	0.353
Middle school	0.368	0.482	0.338	0.473
High school	0.224	0.417	0.207	0.406
College or above	0.186	0.389	0.221	0.415
Working experience (year)	15.37	11.78	12.56	10.72
% urban	0.497	0.5	0.512	0.5
% minority	0.047	0.212	0.05	0.218
Marital status (%)				
Unmarried	0.171	0.377	0.160	0.366
Married	0.804	0.397	0.801	0.399
Separated or widowed	0.024	0.154	0.039	0.194
% poor health	0.074	0.261	0.079	0.270
Obs. #	4795		3420	

Notes: This table provides descriptive statistics for men and women in our sample, with age restricted to between 18 and 60 years old. We report mean average and standard deviation of each variable for men and women, respectively.

TABLE 2 BMI and probability of employment

	Dependent Variable: Employment status							
	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.030*	0.007*	0.030*	0.031*	0.053**	0.014***	0.055**	0.068***
	(0.018)	(0.004)	(0.018)	(0.016)	(0.026)	(0.005)	(0.026)	(0.024)
BMI squared	-0.001		-0.001	-0.001	-0.001**		-0.001**	-0.001***
	(0.000)		(0.000)	(0.000)	(0.001)		(0.001)	(0.001)
Weight status								
Underweight		-0.019				0.035		
		(0.038)				(0.029)		
Normal (Base)								
Overweight		-0.018				-0.107***		
		(0.023)				(0.033)		
Obese		-0.072				-0.315***		
		(0.051)				(0.088)		
Age	0.005	0.005	0.006	0.010**	0.029***	0.028***	0.029***	0.018***
	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.006)
Age squared	-0.000	-0.000	-0.000	-0.000**	-0.000***	-0.000***	-0.000***	-0.000***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Urban *hukou*	0.064***	0.064***	0.057***	0.028*	0.001	0.002	0.000	-0.010
	(0.015)	(0.015)	(0.015)	(0.016)	(0.018)	(0.018)	(0.018)	(0.019)
Han	0.070**	0.070**	0.070**	-0.009	0.075**	0.075**	0.076**	-0.013
	(0.030)	(0.030)	(0.030)	(0.034)	(0.034)	(0.034)	(0.034)	(0.035)
Education level								
Illiterate (Base)								
Primary	0.040	0.039	0.035	0.025	0.033	0.032	0.034	0.018
	(0.032)	(0.032)	(0.032)	(0.029)	(0.037)	(0.036)	(0.037)	(0.030)
Junior	0.101***	0.100***	0.094***	0.076***	0.123***	0.122***	0.126***	0.072**
	(0.029)	(0.029)	(0.029)	(0.027)	(0.033)	(0.033)	(0.033)	(0.028)
Senior	0.159***	0.158***	0.147***	0.112***	0.211***	0.211***	0.214***	0.144***
	(0.031)	(0.031)	(0.031)	(0.029)	(0.035)	(0.035)	(0.036)	(0.032)
College or above	0.281***	0.280***	0.265***	0.227***	0.355***	0.354***	0.355***	0.264***
	(0.031)	(0.031)	(0.031)	(0.030)	(0.035)	(0.035)	(0.036)	(0.033)
Marital status								
Unmarried (Base)								
Married	0.099***	0.099***	0.102***	0.102***	0.017	0.018	0.015	0.015
	(0.023)	(0.023)	(0.023)	(0.021)	(0.029)	(0.029)	(0.029)	(0.027)
Separated or widowed	0.038	0.038	0.035	0.053	-0.023	-0.023	-0.022	-0.003
	(0.049)	(0.049)	(0.049)	(0.046)	(0.050)	(0.050)	(0.050)	(0.048)
Poor health	-0.080***	-0.079***	-0.078***	-0.072***	-0.139***	-0.139***	-0.139***	-0.124***
	(0.026)	(0.026)	(0.026)	(0.024)	(0.031)	(0.031)	(0.031)	(0.029)
Constant	-0.055	0.190	-0.106	-0.045	-0.631**	-0.302**	-0.640**	-0.487*

	(0.224)	(0.121)	(0.225)	(0.208)	(0.309)	(0.150)	(0.310)	(0.290)
Parents' characteristics	No	No	Yes	Yes	No	No	Yes	Yes
County fixed-effects	No	No	No	Yes	No	No	No	Yes
Observations	4,795	4,795	4,795	4,795	3,420	3,420	3,420	3,420
R-squared	0.073	0.073	0.078	0.263	0.094	0.097	0.096	0.301

Notes: This table reports the baseline results for the effects of BMI and its squared term on the probability of being employed for men and women, respectively. The dependent variable is a dummy that equals one for having a job, and zero if otherwise. The first four columns are for the men subsample. Column 1 includes individual covariates only, but does not include parents' characteristics and county fixed effects. Column 2 leaves out the BMI squared term, but a group of weight status dummies. Column 3 includes parents' information in addition to covariates in the specification of Column 1. And Column 4 further includes county level fixed effects. Columns 5-8 are for women with parallel specifications to the first four columns. Robust standard errors are in the parentheses under each coefficient. *** $p < .01$, ** $p < .05$, * $p < .1$.

TABLE 3 BMI and monthly wages

	Dependent Variable: Log value of monthly wages							
	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.004	0.005	-0.005	-0.005	0.396**	-0.002	0.421***	0.419***
	(0.056)	(0.015)	(0.057)	(0.057)	(0.157)	(0.024)	(0.156)	(0.156)
BMI squared	-0.000		0.000	0.000	-0.008**		-0.009***	-0.009***
	(0.001)		(0.001)	(0.001)	(0.003)		(0.003)	(0.003)
Weight status								
Underweight		0.078				-0.400**		
		(0.151)				(0.182)		
Normal (Base)								
Overweight		0.002				-0.034		
		(0.090)				(0.150)		
Obese		-0.171				0.343		
		(0.207)				(0.298)		
Age	0.031	0.031	0.040	0.039	0.068	0.066	0.097**	0.099**
	(0.024)	(0.024)	(0.024)	(0.024)	(0.042)	(0.041)	(0.042)	(0.043)
Age squared	-0.000	-0.000	-0.000	-0.000	-0.001	-0.001	-0.001*	-0.001*
	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Urban *hukou*	0.073	0.072	0.065	0.060	0.045	0.044	0.012	0.012
	(0.057)	(0.057)	(0.065)	(0.065)	(0.098)	(0.097)	(0.118)	(0.119)
Han	0.201	0.199	-0.063	-0.075	0.183	0.177	0.080	0.077
	(0.127)	(0.127)	(0.101)	(0.102)	(0.189)	(0.189)	(0.213)	(0.213)
Experience	0.040***	0.040***	0.044***	0.044***	0.082***	0.082***	0.081***	0.080***
	(0.009)	(0.009)	(0.010)	(0.010)	(0.017)	(0.017)	(0.017)	(0.017)
Experience squared	-0.001***	-0.001***	-0.001***	-0.001***	-0.002***	-0.002***	-0.002***	-0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Education level								
Illiterate (Base)								
Primary	0.354**	0.353**	0.296*	0.266*	0.137	0.138	0.236	0.236
	(0.161)	(0.161)	(0.162)	(0.161)	(0.201)	(0.202)	(0.207)	(0.207)
Junior	0.460***	0.458***	0.365**	0.311**	0.080	0.079	0.214	0.187
	(0.143)	(0.143)	(0.148)	(0.147)	(0.180)	(0.181)	(0.185)	(0.186)
Senior	0.562***	0.561***	0.473***	0.383**	0.231	0.221	0.326	0.278
	(0.145)	(0.145)	(0.150)	(0.149)	(0.192)	(0.193)	(0.201)	(0.202)
College or above	0.950***	0.948***	0.834***	0.681***	0.914***	0.909***	0.985***	0.892***
	(0.151)	(0.151)	(0.157)	(0.164)	(0.188)	(0.188)	(0.199)	(0.213)
Marital status								
Unmarried (Base)								
Married	0.334***	0.336***	0.340***	0.322***	-0.394**	-0.386**	-0.494***	-0.494***
	(0.109)	(0.108)	(0.110)	(0.110)	(0.164)	(0.163)	(0.162)	(0.162)
Separated or widowed	0.023	0.027	0.058	0.054	-0.600**	-0.575**	-0.748***	-0.726***

	(0.226)	(0.227)	(0.220)	(0.220)	(0.242)	(0.240)	(0.256)	(0.256)
Poor health	-0.329***	-0.325**	-0.348***	-0.330***	-0.142	-0.156	-0.106	-0.104
	(0.126)	(0.127)	(0.122)	(0.122)	(0.144)	(0.143)	(0.155)	(0.155)
Constant	5.316***	5.255***	5.486***	5.966***	-0.136	4.623***	-0.843	-0.549
	(0.781)	(0.585)	(0.816)	(0.813)	(1.896)	(0.846)	(1.921)	(1.937)
Parents' characteristics	No	No	Yes	Yes	No	No	Yes	Yes
County fixed-effects	No	No	Yes	Yes	No	No	Yes	Yes
Occupation dummies	No	No	No	Yes	No	No	No	Yes
Observations	3,239	3,239	3,239	3,239	2,284	2,284	2,284	2,284
R-squared	0.077	0.078	0.163	0.172	0.084	0.085	0.197	0.198

Notes: This table reports baseline results for the effects of BMI and its squared term on the wages for men and women conditional on having a job, respectively. The dependent variable is log value of monthly wages. The first four columns are for the male subsample. Column 1 includes individual control variables only. Column 2 follows Column 2 in Table 2 in including a group of weight status dummies. Column 3 includes parents' characteristics and county fixed effects. And Column 4 further includes a group of occupation dummies. The last four columns are for women with parallel specifications. Robust standard errors are in the parentheses under each coefficient. *** p < .01, ** p < .05, * p < .1.

TABLE 4 BMI and labor market attainments for people with different health status

	Employment status				*Log(monthly wages)*			
	Men		Women		Men		Women	
Self-rated health	Good	Poor	Good	Poor	Good	Poor	Good	Poor
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
BMI	0.030*	0.059	0.066**	0.175*	-0.007	-0.646	0.491***	-0.248
	(0.017)	(0.071)	(0.026)	(0.104)	(0.059)	(0.773)	(0.173)	(0.923)
BMI square	-0.001	-0.001	-0.001**	-0.004*	0.000	0.012	-0.010***	0.003
	(0.000)	(0.002)	(0.001)	(0.002)	(0.001)	(0.015)	(0.004)	(0.019)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Parents' characteristics	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	No	No	No	No	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,441	354	3,150	270	3,035	204	2,143	141
R-squared	0.265	0.564	0.301	0.578	0.168	0.648	0.202	0.822

Notes: This table reports results on the effects of BMI and its squared term on labor market attainments for healthier and relatively less healthy men and women, respectively. The odd-numbered columns are for those healthier subsample, while the even-numbered columns are for the less healthy groups. We include individual controls, parents' information and county fixed effects in all specifications, and also occupational dummies in the wage equations. Robust standard errors are in the parentheses under each coefficient. *** p < .01, ** p < .05, * p < .1.

TABLE 5 BMI and monthly wages: by occupation

	Dependent variable: Log value of monthly wages			
	White-collar		Blue-collar	
	Men	Women	Men	Women
	(1)	(2)	(3)	(4)
BMI	-0.002	0.466*	-0.011	0.554**
	(0.087)	(0.262)	(0.081)	(0.220)
BMI squared	-0.000	-0.010*	0.000	-0.012**
	(0.002)	(0.005)	(0.002)	(0.005)
Individual controls	Yes	Yes	Yes	Yes
Parents' characteristics	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes
Observations	1,176	887	1,993	1,360
R-squared	0.320	0.307	0.152	0.208

Notes: This table reports results of BMI and its squared term on wages for white-collar and blue-collar men and women, respectively. The dependent variable is log value of monthly wages. Columns 1 and 2 are for white-collar men and women, while Columns 3 and 4 are for blue-collar men and women. We include individual controls, parents' characteristics, occupation dummies, and county fixed effects in all specifications. Robust standard errors are in the parentheses under each coefficient. *** p < .01, ** p < .05, * p < .1.

TABLE 6 BMI and self-esteem

	Dependent variable: Self-rated score > 3, out of 5					
	Popularity		Get along w/ others		Future confidence	
	Men	Women	Men	Women	Women	Women
	(1)	(2)	(3)	(4)	(5)	(6)
BMI	-0.008	0.009	-0.018	-0.034	0.035	-0.038
	(0.023)	(0.039)	(0.021)	(0.027)	(0.032)	(0.048)
BMI squared	0.000	0.000	0.000	0.001	-0.001	0.001
	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)
Individual controls	Yes	Yes	Yes	Yes	Yes	Yes
Parents' characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Occupation dummies	Yes	Yes	Yes	Yes	Yes	Yes
County fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	4,127	2,885	4,126	2,883	4,127	2,882
R-squared	0.094	0.121	0.076	0.097	0.145	0.137

Notes: This table reports results of BMI and its squared term on proxies for self-esteem level of men and women, respectively. The dependent variables are a dummy that equals one if self-rated score of a certain perspective of self-esteem is higher than three out of five, and zero if otherwise. The first two columns are for the question "How popular do you think you are among others?" Columns 3 and 4 are for the question "How well do you get along with others?" The last two columns are for the question "How confident are you about your future?" We include individual controls, parents' characteristics, occupation dummies, and county fixed effects in each specification. Robust standard errors are in the parentheses under each coefficient. *** p < .01, ** p < .05, * p < .1.