INTERNATIONAL CONFERENCE OF AGRICULTURAL ECONOMISTS
ICAE
29th | Milan Italy 2015
UNIVERSITÀ DEGLI STUDI DI MILANO AUGUST 8 - 14
AGRICULTURE IN AN INTERCONNECTED WORLD

Interaction between Dietary Knowledge and Exercise Knowledge in Leading to Healthier Diet after Hypertension Diagnosis: Evidence from China

By Satoru Shimokawa
Institute of Developing Economies, Japan

Abstract

This paper examines how dietary knowledge and exercise knowledge are interacted each other in leading to healthier diet after hypertension diagnosis. It uses a sharp regression discontinuity design that exploits the exogenous cut-point of systolic blood pressure in the hypertension diagnosis. Using data from China, the results demonstrate that the influence of hypertension diagnosis on improving diet is significant only among people with good dietary knowledge; and within the people with good dietary knowledge, the influence is much larger among the people who do not believe the importance of exercise for health. The findings may imply that an equal emphasis on both diet and exercise might result in reducing the effect of dietary education on improving diet.

Keywords: diet, exercise, hypertension, China
JEL codes: I12, I18, Q18

1. Introduction

There is growing concern in many transition economies about sharp rises in chronic diseases. Particularly, in China, the prevalence of chronic diseases has increased rapidly, and chronic diseases accounted for about 80% of total deaths by 2005 while they accounted for only 65% in 1982 (Bryant, 2003; Wang et al., 2005). Moreover, this trend is predicted to induce substantial health care costs in China. According to the World Health Organization (WHO, 2005), 560 billion U.S. dollars will be foregone during 2000–2015 due to chronic diseases in China – by far the highest health care costs among all of the countries examined in that study. Thus, it is critical for policy makers to prevent and reduce chronic diseases both from a public health perspective and a socioeconomic one.

This paper focuses on the role of dietary and exercise knowledge in preventing or reducing hypertension among adults. In particular, it examines how dietary knowledge and exercise knowledge are interacted each other in leading people to healthier eating and exercise behaviors when people are diagnosed with hypertension. Although several studies have examined the effectiveness of improving dietary/exercise knowledge on changing people's diet/physical activity (Robinson et al. 2003; Datar et al. 2004; Reilly et al. 2006; Taylor et al. 2006; Verstraete et al. 2007; Misra and Kaster 2012), the findings are diverse across studies and tend to be pessimistic (for example, see reviews by Wilks et al. (2011) and Leventer-Roberts et al. (2012)). Also, while previous studies tend to examine the impact of a specific program (one design) by comparing between participants and non-participants, they rarely compare the impacts of different program designs or examine the interaction effect between dietary knowledge and exercise knowledge. In other words, it is sometimes assumed that each of dietary and exercise knowledge has an independent effect and that providing exercise education to people in addition to dietary education would improve exercise behaviors in addition to improving eating behaviors, which is questioned in this paper.

This paper investigates the causal influence of dietary and exercise knowledge on dietary and exercise behaviors after hypertension diagnosis, focusing on the interaction effect between dietary knowledge and exercise knowledge. A key empirical challenge to estimate the causal relationships is potential bias due to omitted variables and reverse causality related to dietary and exercise knowledge and blood pressure (see, for example, Skouteris et al. 2011). First, dietary and exercise knowledge may correlate with some unobserved factors that affect their food consumption and exercise amount (e.g., interests in healthy behaviors), which will cause an omitted variable bias. Second, current blood pressures would depend on the past diet and exercise amount i.e., reverse causality. To overcome both of these problems, we adopt a sharp regression discontinuity (RD) approach which allows us to examine how hypertension diagnosis influences the variation in nutrient intake and exercise patterns among individuals with different initial levels of dietary and exercise knowledge. While our identification strategy is closely related to that in Zhao et al. (2013), our paper is different from the study in two important aspects: (1) we examine the effect of dietary and exercise knowledge and their interaction effect while Zhao et al. (2013) focuses on income level; and (2) our outcome variables include exercise amount in addition to nutrient intake.

We use data from the China Health and Nutrition Survey (CHNS) in 2004, 2006 and 2009. We exploit several features of this data set to control for potential biases. First, the panel nature of the data allows us to condition the outcome variables, nutrient intake and exercise amount, on a diagnosis of hypertension. Second, instead of relying on self-reported hypertension status, we use the blood pressure test results from a physical examination conducted for every individual surveyed in each round of the survey. Finally, since the blood pressure test results are communicated to all survey subjects, the data do not suffer from sample selection bias (see section 2 for more details).

The rest of the paper is organized as follows. Section 2 describes the data and provides background information. Section 3 illustrates the identification and estimation strategies used in the regression-discontinuity design framework. Section 4 presents the estimation results. Second 5 concludes with a discussion of the potential and limitations of dietary/exercise education programs.

2. Background and Data

We use three rounds of panel data from the CHNS in 2004, 2006 and 2009. The CHNS samples are randomly selected from 9 provinces in China applying a sampling with probability proportional to size (PPS) and stratified by income. The CHNS collected three types of data that are critical for this study: (1) nutrition intake and exercise amount; (2) dietary and exercise knowledge; and (3) blood pressure test results. The CHNS also collected data on a wide range of socio-demographic characteristics for each subject.

Our analytical sample consists of individuals who satisfy the following conditions: (i) one is 18 years old or above; (ii) in the current wave, one receives a physical examination (treatment) before a survey about diet and exercise (outcomes); (iii) dietary and exercise knowledge in the previous wave is available; (iv) one has never been diagnosed with hypertension before the physical examination (treatment); and (v) one has never been under treatment of hypertension e.g., anti-hypertension medicine. We need conditions (ii) and (iii) to make sure the timeline of treatment i.e., initial knowledge → treatment → outcomes. Conditions (iv) and (v) are imposed to exclude people who might have expected hypertension diagnosis even before the treatment. As a result, we obtain a panel data set of 4,492 observations. Table 1 presents summary statistics of key characteristics of our analytical sample. Each characteristic is described in the following subsections.

2.1. Nutrition intake and exercise amount

In our sample, the average total daily calorie intake is 2,267.50 kcal. The average calorie share of total fat is 28.8%, that of protein is 11.8%, and that of carbohydrate is 58.4%. Although the total calorie intake is 85.7kcal higher among the control group, calorie shares from the three macronutrients are very similar. The average total calorie intake is moderately higher than 2,100 kcal which is the recommended daily calorie intake for the average male in our analytical sample i.e., one is moderately active and 49 years old with height 161 cm and weight 60 kg. While the calories shares of three macronutrients are within recommended ranges, the calorie share of total fat is relatively high and that of protein is relatively low.

Exercise amount is measured by the summation of the minutes spent for martial arts, gymnastics, dancing, running, swimming, soccer, basketball, tennis, badminton, volleyball, and other (ping poing, Tai Chi, and etc.) during a week. In our sample, the average exercise amount is 29.9 minutes per week, and the average level is similar between the treatment and control groups. WHO (2010) recommends that "adults aged

18–64 should do at least 150 minutes of moderate-intensity aerobic physical activity throughout the week or do at least 75 minutes of vigorous-intensity aerobic physical activity throughout the week or an equivalent combination of moderate- and vigorous-intensity activity”. Thus, even considering measurement errors, exercise amount seems significantly low in our sample, and increasing exercise amount would be desirable.

2.2. Dietary and exercise knowledge

To measure dietary knowledge, we follow the procedure proposed in Shimokawa (2013). We construct a summary index from the answers to the nine diet-related questions in the CHNS (See Appendix 1). In the questions, the subjects choose either ‘agree’, ‘disagree’, or ‘unknown’ for each question. For each of the nine questions, we generate an indicator that takes the value 1 for the correct answer, -1 for the incorrect answer and 0 for choosing ‘unknown’ based on the criteria in WHO (1998). We apply the principal component factor method to the nine indicators in 2004, 2006 and 2009, and construct a summary index of dietary knowledge (DKI) from the first, second and third principal component factors across these nine indicators (i.e., the average of the three factors). We divided our sample at the median of the DKI (= 0.125) into two groups: High Dietary Knowledge group, and Low Dietary Knowledge group.

To measure exercise knowledge, we use the answer to the question “Physical activities are good for one’s health” in the CHNS (See Appendix 1). The subjects choose either ‘agree’, ‘disagree’, or ‘unknown’ for the question. We generate an indicator of good exercise knowledge that takes the value 1 for choosing `agree’, and 0 for choosing other choices. In our sample, 46% of subjects had high exercise knowledge.

2.3. Blood pressure and hypertension diagnosis

To measure hypertension, we use clinical blood pressure levels from individual physical examinations conducted by professionally trained investigators. Systolic and diastolic blood pressure was measured three times for each individual in each round of the survey. Survey personnel informed all the examined subjects about their blood pressures and other results of their physical examinations. When one’s systolic blood pressure is above 140mmHg or one’s diastolic blood pressure is above 90 mmHg, one was diagnosed with hypertension. When the subjects were diagnosed with hypertension, they were also verbally informed the result immediately after the examination. In our analysis, for simplification, we focused on the individuals who were diagnosed with hypertension based on their systolic blood pressure. Accordingly, we excluded individuals who were diagnosed with hypertension only based on diastolic blood pressure (i.e., DBP >= 90 mmHg and SBP < 140 mmHg).

The CHNS sample shows that the prevalence of hypertension increased from 21.7% in 2004 to 26.6% in 2009, and the prevalence of hypertension based on SBP increased from 15.0% to 18.9%. In our analytic sample, we excluded people who ever have diagnosed with hypertension before the physical examination, which cause about 4.1% reduction in the prevalence. As a result, the prevalence of hypertension became 12% in our analytical sample.

It is worth noting that, in our sample, the proportion of individuals with high dietary knowledge is about 50% in each of the treatment and control groups. In contrast, the proportion of individuals with high exercise knowledge is much higher in the treatment group (56%) than in the control group (44%). This may be because exercise knowledge has substantially improved after 2006 while the prevalence of hypertension has also increased during the same period.

3. Estimation Strategy

First, we divide our sample into subsamples based on dietary and exercise knowledge in the previous period. Second, for each subsample, we estimate the causal effect of hypertension diagnosis on nutrient intake and exercise amount in the current period. Lastly, we compare the effect across the subsamples. Figure 1 illustrates the case when we divide our sample into two subsamples based on the level of dietary knowledge in the previous period: a high dietary knowledge (High DK) sample and a low dietary knowledge (Low DK) sample. In each subsample, there is a pair of treatment and control groups, and the impact of hypertension diagnosis on nutrient intake is measured by the difference between the groups. We then test the difference in the impact between the High DK and the Low DK samples.

3.1. Regression Discontinuity Design

To estimate the causal effect, we employ a sharp regression discontinuity (RD) estimation method by exploiting the facts that all survey subjects in the CHNS were informed of their blood pressure test results in each survey round and that an individual is diagnosed with hypertension if one's systolic blood pressure (SBP) is above 140 mmHg. Though diastolic blood pressure above 90 mmHg is also an important indicator of hypertension, patients and physicians often pay closer attention to SBP (Kannel, 2000). Thus, for simplicity this study focuses only on SBP as employed by Zhao et al. (2013).

In our design, the treatment D_i (i.e., hypertension diagnosis) is a deterministic and discontinuous function of a SBP:

$$D_i = \begin{cases} 1 & if\ SBP_i \geq c \\ 0 & if\ SBP_i < c \end{cases} \quad (1)$$

where c is a cut-point and takes the value of 140 mmHg. In the standard parametric econometric specification, the regression to evaluate the treatment effect on the outcome measure Y_i (i.e., nutrient intake and exercise time) is

$$Y_i = \alpha + \rho\, D_i + f(SBP_i) + \eta_i \quad (2)$$

where ρ is the causal effect of hypertension diagnosis, *f*(.) is a flexible continuous function of SBP_i, and η_i is the unexplained variation in Y_i. If equation (2) is linear in D_i and *f*(.) can be correctly specified, the parameter ρ can be estimated using OLS. In OLS estimation, a key issue is the choice of the functional form for *f*(.). Following Lee (2004), we employ quartic polynomials that are preferred for their flexibility in practice.

However, it is still unclear whether the underlying model is in fact linear in D_i and *f*(.). Thus, we estimate the treatment effect noparametrically by employing a local linear regression (LLR) estimation method. In this method, we use only the data close to the cut-off point. Thus the treatment effect is:

$$\rho = Y^{+} - Y^{-} \quad (3)$$

where $Y^{+} = \lim_{SBP \to c^{+}} E[Y_i|SBP_i]$ and $Y^{-} = \lim_{SBP \to c^{-}} E[Y_i|SBP_i]$. As the kernel function for the LLR estimation, we use a triangular kernel. We chose the bandwidth for the kernel function following the cross-validation procedure proposed in Imbens and Lemieux (2008).

3.2. Preliminary Checks on the Regression Discontinuity Design

There are two assumptions required for an appropriate RD design: (i) the individuals cannot precisely control the value of SBP; and (ii) the individual characteristics right above or below the cut-point do not differ systematically.

In our setting, the first assumption holds because one cannot precisely control one's SBP at a particular time when one received a physical examination in the survey. Although one can influence blood pressure by taking antihypertension drugs, we excluded people who are taking such drugs. We also excluded the people who have ever been diagnosed with hypertension before the examination.

To examine whether the second assumption holds, we start checking the distribution of key observable socioeconomic factors by blood pressure. Figure 2 (a) shows that these factors are distributed continuously around the cut-point of 140 mmHg. To check for unobserved factors, we examine the distribution of the assignment variable itself as suggested in Lee and Lemieux (2010). Figure 2 (b) shows that the kernel density of SBP is approximately normally distributed, without a notable change in its distribution around the cut-point. Since there is no systematic difference between the two samples to the left and to the right of the cut-point, if we observe changes in the outcomes variables of interest, they are likely to be due to the treatment, i.e. hypertension diagnosis.

4. Results

We used local linear regressions (LLR) to estimate the impact of hypertension diagnosis on nutrient intakes (total calorie, fat, carbohydrate, and protein intake) and exercise amount. We use individual systolic blood pressure (SBP) as our assignment variable with the cut-point of 140mmHg, and daily nutrient intake and weekly exercise amount after the treatment (within the same wave) as our outcome variable. Using the cross-validation approach proposed by Imbens and Kalyanaraman (2009), we estimate the consistent optimal bandwidth around the cut-point by minimizing the mean integrated squared error. To examine the robustness of the LLR estimates, we estimate the impact using 50% and 200% of the optimal bandwidth. In addition, we estimate the parametric specification with quartic polynomials with and without other covariates. The numeric results from the LLR and parametric estimations are presented in tables 2 and 3.

4.1. Main Results

We begin with a graphical representation of the observed relationship between systolic blood pressure levels and daily nutrition intake and weekly exercise amount. Figure 3 present averages (circles) and local linear smoothers (solid curves) of the daily intake of total calorie and three macronutrients (fat, carbohydrate, and protein), and weekly exercise amount plotted against systolic blood pressure. These figures are shown for the whole sample as well as separately for a high dietary knowledge (High DK) and a low dietary knowledge (Low DK) group, or for a high exercise knowledge (High ExK) and a low exercise knowledge (Low ExK) group.

Figure 3 demonstrates four important relationships. First, there is a clear discontinuity in total calorie intake for both the High DK group and the Low DK group, and hypertension diagnosis had a negative impact in both groups. Second, in the High DK group, the negative impact on total calorie intake is largely attributable to the negative impact on fat intake. Third, in the Low DK group, the negative impact on total calorie intake is largely attributable to the negative impact on carbohydrate intake. Lastly, there is a clear discontinuity in exercise amount only for the High ExK group while not for the Low ExK group.

The numeric LLR estimates in Table 2 support the graphical findings for nutrient intake, while do not support those for exercise amount. The estimates show that hypertension diagnosis reduced total calorie intake by 163.8kcal in the High DK group (a 7.4% decline from the mean level) and by 137.8kcal in the Low DK group (a 6.1% decline from the mean level). Fat intake is significantly reduced only in the High DK group (-13.6g), while carbohydrate and protein intakes are significantly reduced only in the Low DK group (-36.1g and -10.4g, respectively). In contrast, while the estimated effect on exercise amount is positive, it is not statistically significant even in the High ExK group.

In addition, we examine the interaction between dietary knowledge and exercise knowledge by defining four distinct initial knowledge levels: High DK and High ExK, High DK and Low ExK, Low DK and High ExK, and Low DK and Low ExK. The LLR estimates show that hypertension diagnosis significantly reduced total calorie intake only among individuals with High DK and Low ExK (-258.2kcal = a 11.7% decline from the mean level), and the decline is attributable to the negative impact on fat intake (-29.0g = -270kcal). It is also worth noting that, even within the high DK group, if individuals also have high ExK, hypertension diagnosis could have only insignificant impact on all nutrient intakes. Similarly, within the low DK group, the impact on nutrient intake is larger when people do not believe the importance of exercise i.e., Low ExK. Among individuals with Low DK and Low ExK, hypertension diagnosis reduced total calorie intake by 149.9kcal while the effect is not statistically significant, and the negative effect is mainly attributable to the significant effect on carbohydrate intake (-50.7g).

4.2. Robustness Checks

The above LLR estimates can be sensitive to the choice of bandwidth and specifications. Thus, we first examine the estimates by using different bandwidths (50% and 200% of the optimal bandwidth). Our local linear regression results on nutrient intake and exercise amount seem robust to the choice of different bandwidths. Either an increase or a decrease in the bandwidth does not lead to significantly different results.

Second, we have also estimated parametric regressions as presented in Tables 2 and 3. Including additional controls tends to increase the magnitude of the estimates for the High DK group and reduce the magnitude of the estimates for the Low DK group, which suggests that some of these controls may be correlated with the treatment variable, causing a upward bias in the estimates for the High DK group and a downward bias in the estimates for the Low DK group. This is a concern for the parametric results. Since the parametric estimation is based on all the data, including observations “far away” from the cut-point, a poor approximation of the function $f(.)$ in equation (2) may result in the violation of the RD assumptions and, consequently, lead to biased estimates.

Although our parametric estimates tend to be larger in magnitude than our LLR estimates, it is hard to compare the magnitude between these two methods. Parametric estimates may be biased if a polynomial is a poor approximation of the function $f(.)$, while the LLR estimates may also be biased if the model is non-linear even within a close neighborhood of the cut-point. However, considering the robustness of the LLR estimates to smaller bandwidths and the sensitivity of the parametric estimates to inclusion of control variables, the LLR estimates seem more likely to be consistent than the parametric estimates.

5. Discussions

This paper examined how dietary and exercise knowledge affect eating and exercise behaviors when people are diagnosed with hypertension. Using data from China, we find that, following hypertension diagnosis, nutrient intake was improved significantly only among people with good dietary knowledge; and within the people with good dietary knowledge, the improvement was much larger among the people who do not believe the importance of exercise for health compared to people who believe it. The same tendency was also observed among people with low dietary knowledge i.e., the impact on nutrient intake was larger among people who do not believe the importance of exercise for health. In contrast, we found no significant impact of hypertension diagnosis on weekly exercise amount.

These findings imply that, at least in our Chinese sample, healthy diet and exercise are substitute rather than complement, and people choose one of them instead of improving both. Also, the results may imply a potential risk of combining dietary and exercise education. For example, an equal emphasis on both healthy diet and exercise might result in reducing the effect of dietary education on improving diet while exercise education has insignificant effect on exercise amount.

References

Bryant, C.D., 2003. Handbook of Death and Dying. The Sage Publication, Thousand Oaks, CA.

Datar, A., & Sturm, R. (2004). Physical education in elementary school and body mass index: Evidence from the early childhood longitudinal study. *American Journal of Public Health*, *94*(9), 1501-1506.

Imbens, G., Lemieux, T., 2008. Regression discontinuity designs: a guide to practice. Journal of Econometrics 142 (2), 615–635.

Kannel, W.B., 2000. Elevated systolic blood pressure as a cardiovascular risk factor. American Journal of Cardiology 85 (2), 251–255.

Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. Journal of Economic Literature 48 (2), 281–355.

Leventer-Roberts, M., A. Patel, and L. Trasande. 2012. "Is severity of obesity associated with diagnosis or health education practices?" International Journal of Obesity, 36: 1571-1577.

McFerran, Brent and Anirban Mukhopadhyay. 2013. "Lay Theories of Obesity Predict Actual Body Mass." *Psychological Science*, forthcoming.

Ma, J. (2005). Fitness plan calls for all villages to get sports facilities. *South China Morning Post*.

Muntner, P., Gu, D., Wildman, R. P., Chen, J., Qan, W., Whelton, P. K., & He, J. (2005, September). Prevalence of Physical Activity Among Chinese Adults: Results From the International Collaborative Study of Cardiovascular Disease in Asia. *Am J Public Health, 95*(9), pp. 1631-1636. doi:10.2105/AJPH.2004.044743

Lee, S., R. Misra, and E. Kaster. 2012. "Active Intervention Program Using Dietary Education and Exercise Training for Reducing Obesity in Mexican American Male Children", the Health Educator, 44(1): 2-13.

Reilly JJ, Kelly L, Montgomery C, Williamson A, Fisher A, McColl JH, Lo Conte R, Paton JY, Grant S. Physical activity to prevent obesity in young children: cluster randomised controlled trial. *BMJ* 2006; **333**: 1041–1045.

Robinson TN, Killen JD, Kraemer HC, Wilson DM, Matheson DM, Haskell WL, Pruitt LA, Powell TM, Owens AS, Thompson NS, Flint-Moore NM, Davis GJ, Emig

KA, Brown RT, Rochon J, Green S, Varady A. Dance and reducing television viewing to prevent weight gain in African–American girls: the Stanford GEMS pilot study. *Ethn Dis* 2003; **13**(Suppl.): 65–77.

H. Skouteris, M. McCabe, B. Swinburn, V. Newgreen, P. Sacher, and P. Chadwick. 2011. “Parental influence and obesity prevention in pre-schoolers: a systematic review of interventions” Obesity Reviews, 12: 315-328.

Taylor RW, McAuley KA, Williams SM, Barbezat W, Nielsen G, Mann JI. Reducing weight gain in children through enhancing physical activity and nutrition: the APPLE project. *Int J Pediatr Obes* 2006; **1**: 146–152.

Verstraete SJ, Cardon GM, De Clercq DL, De Bourdeaudhuij IM. A comprehensive physical activity promotion programme at elementary school: the effects on physical activity, physical fitness and psychosocial correlates of physical activity. *Public Health Nutr* 2007; **10**: 477–484.

Wang, L., Kong, L., Wu, F., Bai, Y., Burton, R., 2005. Preventing chronic diseases in China. Lancet 366 (9499), 1821–1824.

Wilks, D.C., H. Besson, A.K. Lindroos, and U. Ekelund. 2011. “Objectively measured physical activity and obesity prevention in children, adolescents and adults: a systematic review of prospective studies.” Obesity Review, 12: e119-e129.

World Health Organization, 1998. Development of Food-based Dietary Guidelines for the Asian Region. Intercountry Workshop on Implementation of Research Agenda of International Conference on Nutrition, SEARO/New Delhi, India.

World Health Organization, 2005. Preventing Chronic Diseases: A Vital Investment. WHO Global Report.

World Health Organization, 2010. Global Recommendations on Physical Activity for Health.

Zhao, Meng, and Yoshifumi Konishi, and Paul Glewwe. 2013. “Does Information on Health Status Lead to a Healthier Lifestyle? Evidence from China on the Effect of Hypertension Diagnosis on Food Consumption.” *Journal of Health Economics*, 32(2): 367-385.

Table 1: Summary Statistics of our analytical sample

	All n = 4,492 Mean	(SD)	Treatment n = 565 Mean	(SD)	Control n = 3,927 Mean	(SD)
Nutrient intake and exercise outcome						
Initial total calorie intake (kcal)	2267.50	(658.92)	2192.18	(656.70)	2277.88	(658.63)
Initial fat intake (g)	72.52	(36.37)	71.52	(36.79)	72.66	(36.31)
Initial protein intake (g)	66.93	(24.12)	64.76	(23.17)	67.23	(24.23)
Initial carbohydrate intake (g)	331.15	(116.63)	316.45	(114.30)	333.17	(116.82)
Initial exercise amount (min/week)	29.89	(156.76)	28.17	(129.76)	30.12	(160.01)
Dietary and exercise knowledge						
Prop of High Dietary Knowledge	0.50	(0.50)	0.50	(0.50)	0.50	(0.50)
Prop of High Exercise Knowledge	0.46	(0.50)	0.56	(0.50)	0.44	(0.50)
Hypertension and blood pressure						
Hypertension diagnosis indicator	0.12	(0.32)	-	-	-	-
Systol blood pressure (mmHg)	122.43	(14.05)	150.56	(10.71)	118.64	(9.36)
Other individual characteristics						
Age (years)	48.62	(13.45)	58.14	(12.43)	47.34	(13.06)
Female indicator	0.52	(0.50)	0.54	(0.50)	0.51	(0.50)
Household income per member (yuan)	8370.09	(12366.13)	8850.60	(9246.05)	8305.42	(12727.24)
High education indicator	0.60	(0.49)	0.48	(0.50)	0.61	(0.49)

Table 2: Numeric Estimates of the Effect of Hypertension Diagnosis on Nutrient Intake and Exercise Amount by Initial Dietary/Exercise Knowledge Level

		All	High DK	Low DK
	Column #	(1)	(2)	(3)
Observations		4,492	2,457	2,257
Total calorie intake (kcal)				
Non-parametric (bw = 18.6)		-140.5**	-163.8*	-137.8
Parametric	No controls	-204.2***	-199.0**	-231.2**
	With controls	-181.1***	-201.9**	-189.5**
Fat intake (g)				
Non-parametric (bw = 7.5)		-6.9	-13.6^	1.9
Parametric	No controls	-6.9*	-9.7*	-3.8
	With controls	-6.2	-10.8**	-1.5
Carbohydrate intake (g)				
Non-parametric (bw = 11)		-25.5**	-22.6	-36.1**
Parametric	No controls	-29.6***	-25.3*	-40.1***
	With controls	-28.4***	-25.2*	-38.8***
Protein intake (g)				
Non-parametric (bw = 7)		-8.0**	-5.1	-10.4*
Parametric	No controls	-5.9***	-4.3	-7.8**
	With controls	-4.8**	-4.5^	-5.9*
		All N = 2876	High ExK N = 1297	Low ExK N = 1490
Exercise amount (min/week)				
Non-parametric (bw = 12.6)		24.1	11.0	5.1
Parametric	No controls	-9.1	-21.6	-5.2
	With controls	3.8	1.1	-2.6

Note: (1) *** = 1%, ** = 5%, * = 10%, and ^ = 15%.
(2) DK = dietary knowledge, ExK = exercise knowledge, and bw = optimal bandwidth.

Table 3: Numeric Estimates of the Effect of Hypertension Diagnosis on Nutrient Intake and Exercise Amount by Initial Dietary × Exercise Knowledge Level

		High DK High ExK	High DK Low ExK	Low DK High ExK	Low Dk Low ExK
	Column #	(1)	(2)	(3)	(4)
Observations		674	765	607	691
Total calorie intake (kcal)					
Non-parametric (bw = 18.6)		20.9	-258.2*	-51.8	-149.9
Parametric	No controls	51.9	-293.0**	-241.7	-56.6
	With controls	-11.8	-288.4*	-79.1	-2.1
Fat intake (g)					
Non-parametric (bw = 7.5)		-14.5	-29.0*	21.2**	-5.7
Parametric	No controls	-3.9	-16.4*	-7.5	1.2
	With controls	-4.3	-19.9**	-0.4	3.0
Carbohydrate intake (g)					
Non-parametric (bw = 11)		12.5	-23.9	-19.9	-50.7^
Parametric	No controls	15.9	-28.6	-28.6	-15.9
	With controls	3.1	-16.8	-14.4	-12.1
Protein intake (g)					
Non-parametric (bw = 7)		1.2	-2.7	2.6	-3.7
Parametric	No controls	1.3	-1.5	-8.1	-2.0
	With controls	0.2	-3.7	-2.2	0.6
Exercise amount (min/week)					
Non-parametric (bw = 12.6)		127.2	39.0	-49.9	-15.4
Parametric	No controls	13.0	-29.9	-56.7^	44.6
	With controls	44.6	-75.9^	-41.2	69.2**

Note: (1) *** = 1%, ** = 5%, * = 10%, and ^ = 15%.
(2) DK = dietary knowledge, ExK = exercise knowledge, and bw = optimal bandwidth.

Figure 1: An Example of Our Analytical Strategy

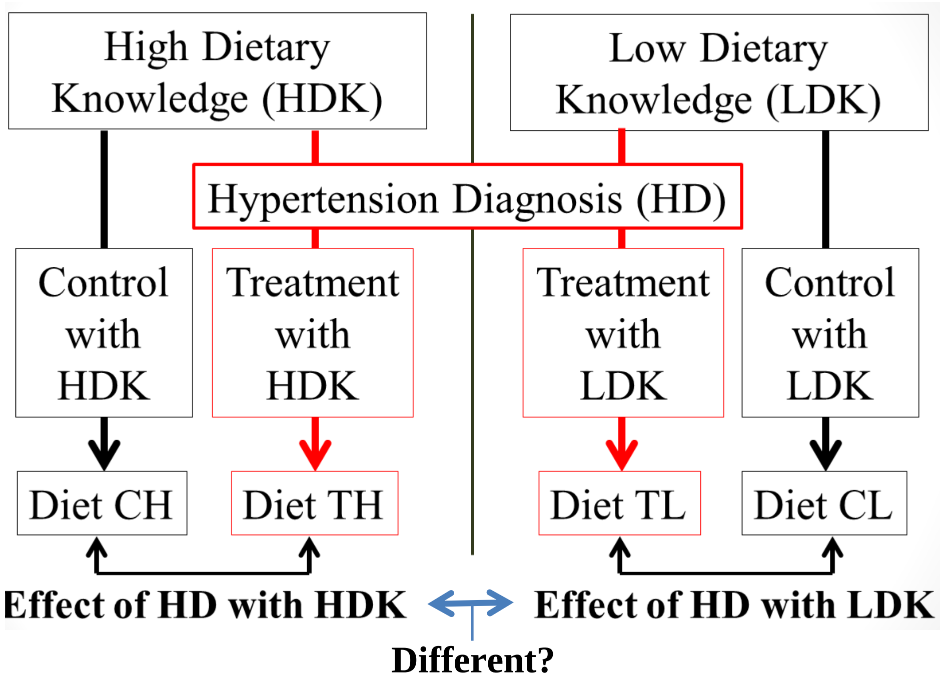
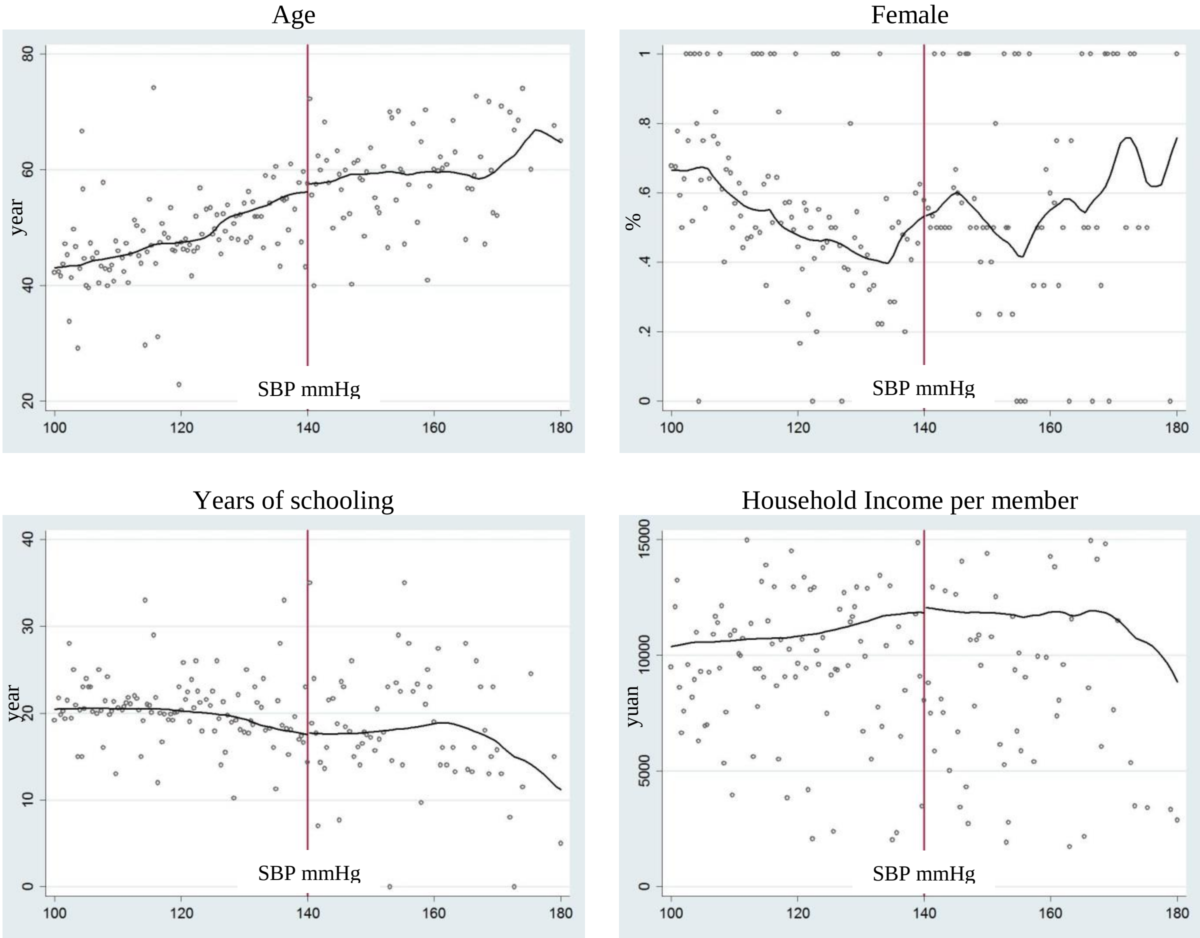
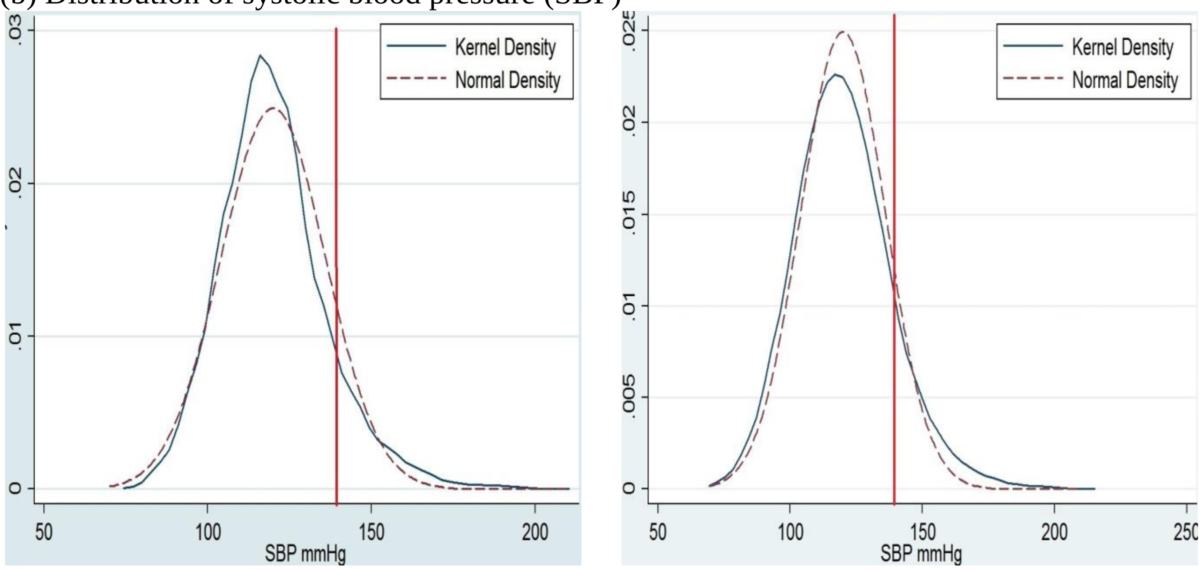


Figure 2: Difference in individual characteristics right above or below the cut-point

(a) Local mean smoothing of key socioeconomic variables by systolic blood pressure



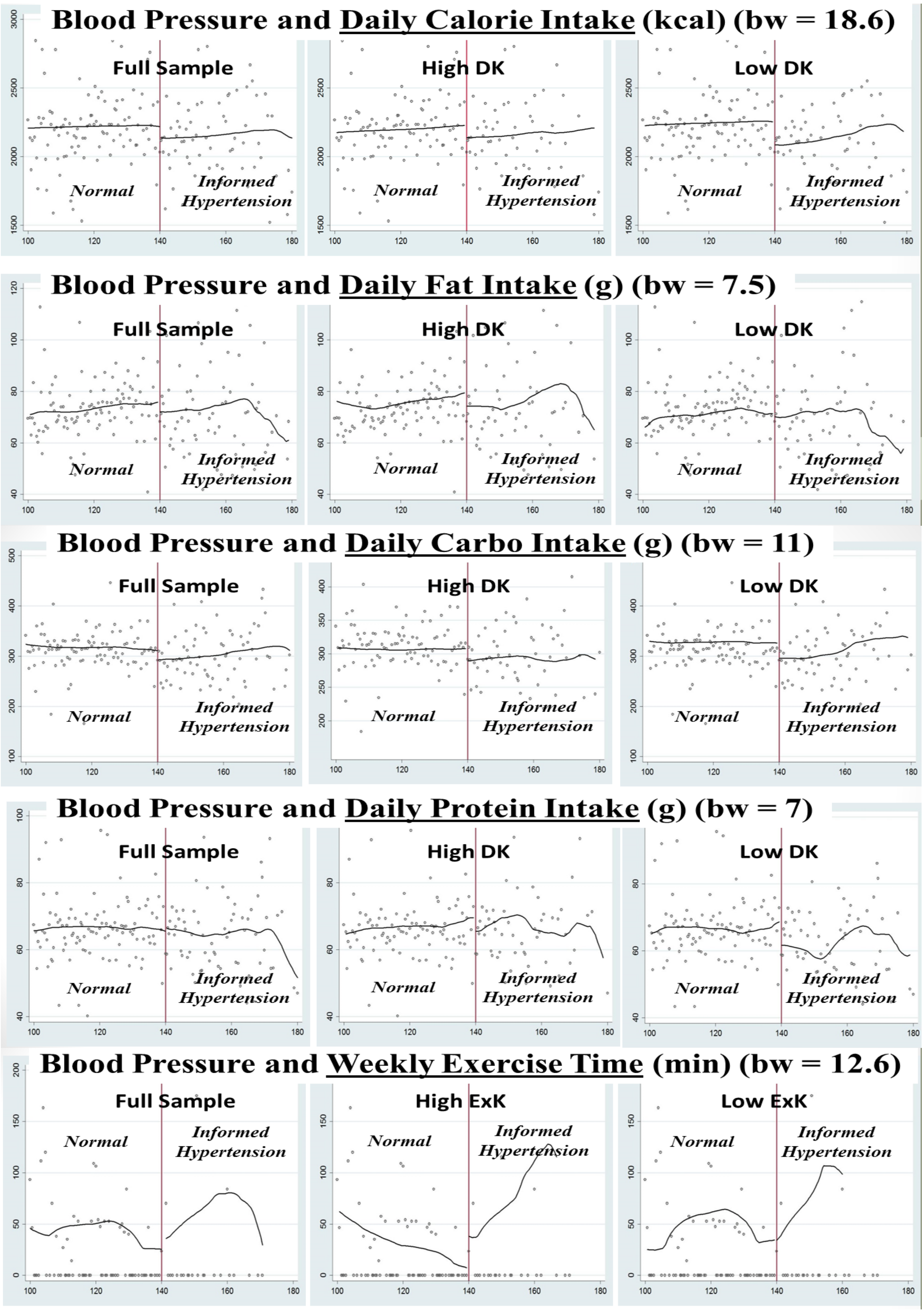
(b) Distribution of systolic blood pressure (SBP)



Bandwidth = 5 mmHg

Bandwidth = 10 mmHg

Figure 3: Non-parametric Estimates of the Effect of Hypertension Diagnosis on Nutrient Intake and Exercise Amount by Initial Knowledge Level



Appendix 1: Questions used to measure dietary and exercise knowledge

	Statement Do you strongly agree, somewhat agree, somewhat disagree or strongly disagree with this statement? * Please note that the question is not asking about your actual habits.	1 strongly disagree 2 disagree 3 neutral 4 agree 5 strongly agree 9 unknown
1	Choosing a diet with a lot of fresh fruits and vegetables is good for one's health.	
2	Eating a lot of sugar is good for one's health.	
3	Eating a variety of foods is good for one's health.	
4	Choosing a diet high in fat is good for one's health.	
5	Choosing a diet with a lot of staple foods [rice, wheat and related products] is not good for one's health.	
6	Consuming a lot of animal products daily (fish, poultry, eggs and lean meat) is good for one's health.	
7	Reducing the amount of fatty meat and animal fat in the diet is good for one's health.	
8	Consuming milk and dairy products is good for one's health.	
9	Consuming beans and bean products is good for one's health.	
10	Physical activities are good for one's health.	