Stratum Differences in Health Behaviour of Elderly Group—
An empirical research based on harmonized CHARLS

Jie Hua1*
1Department of Sociology, School of Humanities, Southeast University, 210000, Nanjing, China
*213160515@seu.edu.cn

Abstract. Based on the harmonized CHARLS data, this paper discussed the impact of social class on the health behaviour of the elderly. Through negative binomial panel regression and logit panel regression with fixed effect, the study found that older people with higher social strata pay more attention to maintaining their health and thus tend to avoid health risk behaviours. Among them, education has a specially significant impact on the choice of health behaviour, and elder people with higher education have a healthier behaviour. In addition, retirement also has a significant impact on the health of the elderly, mainly reflected in the possibility of their exercise habits disappearing after retirement.

1. Introduction
Health has always been a research issue in the field of social science. The World Health Organization (WHO) believes that health is not only free from disease or injury, but also a sound state of physical, psychological and social well-being. Thus, health is also one of the core agenda of social development. In the post-medical era of developed society, more and more environmental factors and social factors which affect health have been paid attention to.

According to the division of American health century, it was regarded as the environmental era before 1925, the medical era from 1925 to 1950, and then the era of lifestyle. In the era of lifestyle, the topics include the discussion of diet balance, fitness exercise, tobacco and alcohol culture, which can be said to cover a variety of behavioural choices affecting physical health. Since the 1960s, with smallpox and other persistent diseases being eliminated by large-scale vaccine development, the mortality rate of infectious diseases has been limited, and people have focused on health regulation beyond the pathological level [1].

The heterogeneity of social groups has always been one of the foundations of sociological research, that is, there are economic, cultural, political and health differences among people. Stratification is a term borrowed from geography by sociologists. In the field of sociology, social stratification is used to describe the systematic structure of class differences and inequalities. Early studies have proved that social and economic status differences can lead to individual health and life expectancy [2]. Since then, in the discussion of social stratification, health issues have also been concerned.

Because the elderly are in the later stage of their life course, their health problems are particularly important [3]. Therefore, the research questions of this paper are: how the social class attributes of the elderly group affect their health behaviour choice at present, and on this basis, the reasons for this phenomenon are discussed.
2. Theory and Literature Review
We first make a brief review of the essential researches and mainstream theories.

The explanation of health inequality mainly concentrates on two different theories, namely social causality theory and health choice theory. These two theories mainly come from the theoretical summary of health stratification by Western sociologists. Chinese scholars test the theory practically and compare the explanatory power of which theory is stronger in China's national conditions. Using tracking data, such as growth curve model, multiple linear regression model and so on, to verify and compare. After comprehensive discussion in many aspects, a convincing conclusion is that social causality theory in China is more explanatory than health choice theory [4].

Although the positive correlation between socioeconomic status and health has been basically established, the question of how this relationship changes with age is far from unanimous. Some studies have found that health disparities among different socio-economic status groups widen gradually before middle age and the early years of old age, and shrink continuously in old age, some studies call it convergence hypothesis [6]. As for the reasons for the shrinking of health disparities in old age, it may be that in old age, the differences in psychosocial risk factors (such as lack of social relations, etc.) faced by different social and economic status groups may gradually narrow or even disappear. As for the reasons for the shrinking of health disparities in old age, on one hand, it may be that in old age, the differences in psychosocial risk factors (such as lack of social relations and social support, loss of sense of control, etc.) faced by different social and economic status groups are gradually narrowing or even disappearing [7]. On the other hand, it may be that the determinant effect of biological factors on health is increasing gradually, even surpassing that of socioeconomic factors.

However, some scholars have put forward opposite opinion. Cumulative advantage hypothesis holds that the impact of socioeconomic status on health will continue to accumulate throughout the whole life. The health disparities among people with different socioeconomic status are increasing rather than decreasing with age, and the health inequality in old age is greater than that in middle age [9].

In a word, there is no consistent conclusion about the relationship between social and economic status and health in the elderly stage, and most of these conclusions are based on the experience of developed countries. Therefore, it is urgent to re-examine this issue in a very different socio-economic context. Based on the above discussion, we will discuss further on the basis of social determinism.

3. Datasets and Variables
In this chapter, we mainly introduce the datasets and selected variables in the research.

3.1 Datasets
The dataset used in this research is harmonized CHARLS.

China Health and Retirement Longitudinal Survey is a large-scale interdisciplinary survey project sponsored by the National Development Institute of Peking University and co-executed by the Chinese Social Science Research Center of Peking University and the League Committee of Peking University. It is a major project funded by the National Natural Science Foundation of China. It aims to collect a set of projects representing China aged 45 and older.

CHARLS National Baseline Survey was conducted in 2011 and followed up every two years (2011, 2013, 2015), so the dataset used in this research is a typical panel dataset.

3.2 Variables
The variables selected in this paper can be divided into two parts, one is health behaviour, the other is social class attributes.

After data cleaning, the cases with missing values were deleted, and the final effective and complete dataset consists of 6362 cases.

To measure the health behaviour, we choose smoke, drink and physical exercise. Smoke and drink are enumeration variables, which depicts the frequency of smoking and drinking. Physical exercise is a binary variable, indicating whether the respondent participated in physical exercise weekly.
To measure the social class, we choose income and education. Income is measured by the total income after tax in the past year, while we also make a logarithmic transformation of it to satisfy the normal distribution assumption. Education is a binary variable that indicates whether the respondent has received higher education or not.

Besides the variables mentioned above, we also include some demographic information, such as gender and age. Gender is a binary variable that 0 represents female while 1 indicates male. Age is a continuous variable, and since we mainly focus on the elderly group, people under 45 years old are excluded. Social activity is a binary variable that indicates the respondent whether participated in any social activities monthly.

Table 1. Descriptive Statistics and Basic Information of Variables

| Variable   | Mean     | Standard Deviation | Remark                      |
|------------|----------|--------------------|-----------------------------|
| Drink      | 1.172    | 2.307              | Frequency of drinking weekly|
| Smoke      | 4.077    | 9.162              | Frequency of smoking daily  |
| Exercise   | /        | /                  | 1 represents participation  |
| Income     | 9.352    | 1.330              | Logarithm of yearly income  |
| Education  | /        | /                  | 1 represents higher education|
| Gender     | /        | /                  | 1 represents male           |
| Age        | 57.857   | 9.569              | Minimum is 45, Maximum is 97|
| Social Activity | /     | /                  | 1 represents participation  |

4. Data Analysis

This chapter mainly consists of two parts, one focuses on the effect of social economic status and the other part discusses the influence of retirement.

4.1 Health Behaviour and Socioeconomic Status

To analyse the panel data, the model used is also determined by the type of dependent variable as cross-section data.

It is noted that the frequency of smoking and drinking is a counting variable, so a suitable choice is Poisson regression or negative binomial regression. To make a trade-off between Poisson regression and negative binomial regression, we will further look into whether there exists over dispersion of dependent variable [10].

The ratio of variance and mean of drink is 4.54 and the that of smoke is 20.59, the over dispersion is obvious, so a more proper strategy here is to use the panel negative binomial regression.

Since this is a panel data, we must also consider the trade-off between the random effect model and the fixed effect model. For the negative binomial panel regression with random effects, it assumes that the individual heterogeneity obeys the Beta distribution, therefore, the individual heterogeneity can be integrated, and then the maximum likelihood estimation is made. For the negative binomial panel regression with fixed effects, the conditional likelihood function is considered under given sufficient statistics of individual heterogeneity $n_i = \sum_{t=1}^{T} y_{it}$, and then the conditional maximum likelihood estimation is performed.

Here we use the commonly used Hausman Test to determine whether we should choose random effect model or not and LR Test (Likelihood-Ratio Test) to determine whether we should use pooled negative binomial regression.

According to Hausman Test and LR Test, finally the Logistic panel regression with fixed effect is chosen to be implemented for variable Exercise (accuracy of numerical integration is in an acceptable range, all the relative difference of all coefficients is less than $10^{-4}$). Negative binomial panel regression with random effect is selected for variable drink, and pooled negative binomial panel regression is chosen for variable smoke.
Table 2. Results of Panel Regression

|                | Exercise       | Drink         | Smoke         |
|----------------|----------------|---------------|---------------|
| Age            | 0.002 (0.003)  | 0.002 (0.003) | 0.005 (0.006) |
| Gender         | 2.808**(0.097) | 2.831**(0.097) | 3.667**(0.138) |
| Income         | 0.031 (0.014)  | 0.132**(0.014) | -0.078**(0.039) |
| Education      | -0.153 (0.098) | -0.151**(0.093) | -0.771**(0.164) |
| Social Activity| 0.119**(0.030) | 0.168**(0.032) | 0.298**(0.076) |
| Constant       | /              | 2.781**(0.275) | 2.645**(0.490) |

(1)**P<0.10, **P<0.05, ***P<0.01 (2-tailed Test); (2) Numbers in parentheses are standard errors, which is calculated by bootstrap.

First, an interesting phenomenon is that age is not significant in these three models, which means that there is no age difference in healthy behavior as a habit.

As far as gender is concerned, men are more likely to have the habit of smoking and drinking, but they also pay more attention to individual exercise.

In terms of income, higher-income groups pay more attention to exercise and smoke less frequently. The relationship between income and alcohol consumption is not highly significant, but the higher the income, the higher the frequency of alcohol consumption will increase accordingly. One possible explanation is that groups with higher incomes may have more social interaction and thus need to drink. In addition, there are different types of alcohol, such as high-quality red wine, which often represents the taste and relationship of the upper class.

Education has no significant effect on physical exercise behavior, while the group with higher education is more likely to smoke and drink less.

The influence of social activities on healthy behavior is significant, with both positive and negative effects. In terms of exercise, socially active groups tend to exercise more. In terms of smoking and drinking, groups with active social participation also have higher frequency of smoking and drinking.

4.2 Effect of Retirement

Although age was not significant in regression analysis, it could not be concluded that age had no effect on healthy behavior. Here we are interested in the impact of retirement on healthy behavior, because there will be more social interaction before retirement, and the corresponding exercise time may be less.

A suitable choice is regression discontinuity. The basic idea of regression discontinuity is that when a continuous variable exceeds a certain threshold, it will produce a certain jump, thus overcoming the problem of treatment effect and making causal inference [11]. Generally, regression discontinuity can be divided into two categories, one is sharp regression discontinuity and the other one is fuzzy regression discontinuity. The feature of sharp regression discontinuity is that at x=c, the probability of individuals being treated jumps from 0 to 1. The feature of fuzzy regression discontinuity is that at x = c, the probability of individuals being treated jumps from a to b (0<a<b<1).

Here we only use the cross-sectional data of CHARLS 2015. The statutory retirement age in China is 60 years old for men and 55 years old for women, so we divide the samples into two groups, one consists of male and the other one consists of female. Since it may take some time to reflect the impact of retirement, we implement the fuzzy regression discontinuity with covariates income, education, and social activity. The local Wald estimate is as below:

Table 3. Local Wald Estimate of Retirement

|                | Male            | Female          |
|----------------|-----------------|-----------------|
| Exercise       | -0.228**(0.093) | -0.187**(0.085) |
| Drink          | 0.093(0.568)    | 1.259(0.925)    |
| Smoke          | 0.779(0.613)    | -2.716(2.524)   |

*P<0.10, **P<0.05, ***P<0.01(2-tailed Test)

The optimal bandwidth in the model is set to 100. Normally, the optimal bandwidth values of 50 and 200 need to be reported, but we leave out it for that the results are highly consistent with the above.
A surprising finding is that there is no significant difference in drinking and smoking behavior before and after retirement. This proves that risky health behavior, as a long-term cumulative habit, does not change significantly influenced by retirement. And no matter men or women, behavior of physical exercise will reduce significantly after retirement. In addition to taking into account the decline in physical fitness, another point may be that many elderly people take on the responsibility of caring for their grandchildren after retirement, so that they do not intentionally exercise.

5. Conclusion and Discussion

Through the above data analysis, we may draw some conclusions.

Generally speaking, there is undoubtedly a stratification in health behavior choice. Relatively speaking, the older groups of higher strata tend to avoid health risks and pay more attention to maintaining their health. Among them, education plays an important role. As far as education is concerned about the cultivation and shaping power of individuals, the higher the level of education, the higher the awareness of the importance of health. Social interaction has both positive and negative effects, which means that sometimes people are forced to take risky health behaviors, such as drinking, to maintain social networks.

After retirement, the old people still have the original risk health behavior, and exercise habits may disappear, thus their health risks may further increase.

In order to better care for the health of the elderly, on the one hand, more basic health policies need to be formulated to take care of the physical health of the middle and lower class, on the other hand, it is necessary to promote some sports for the elderly.

References

[1] Cockerham, W. C. (2000). *Health lifestyles in Russia*. Social Science & Medicine, 51(9), 1313-1324.
[2] Krieger, N., Rowley, D. L., Herman, A. A., Avery, B., & Phillips, M. T. (1993). *Racism, sexism, and social class: implications for studies of health, disease, and well-being*. American Journal of Preventive Medicine, 9(6 Suppl), 82.
[3] Luo, Y., Hawkley, L. C., Waite, L. J., & Cacioppo, J. T. (2012). *Loneliness, health, and mortality in old age: a national longitudinal study*. Social Science & Medicine, 74(6), 907-914.
[4] F.Q. Wang. (2012). *Socio-economic status, lifestyle and health inequalities*. Society (2), 125-143.
[5] Beckett, M. (2000). *Converging health inequalities in later life--an artifact of mortality selection*. Journal of Health & Social Behavior, 41(1), 106-119.
[6] H. Xu, Y. Xie. (2017). Socioeconomic inequalities in health in china: a reassessment with data from the 2010–2012 china family panel studies. Social Indicators Research, 132(1), 219-239.
[7] House, J. S., Lepkowski, J. M., Kinney, A. M., Mero, R. P., & Herzog, R. C. K. R. (1994). *The social stratification of aging and health*. Journal of Health and Social Behavior, 35(3), 213-234.
[8] Mirowsky, J., & Ross, C. E. (2005). *Education, cumulative advantage, and health*. Ageing International, 30(1), 27-62.
[9] Dupre, M. E. (2008). *Educational differences in health risks and illness over the life course: a test of cumulative disadvantage theory*. Social Science Research, 37(4), 1253-1266.
[10] Gardner, W, Mulvey, E. P., & Shaw, E. C. (1995). *Regression analyses of counts and rates: poisson, overdispersed poisson, and negative binomial models*. Psychological Bulletin, 118(3), 392-404.
[11] Imbens, G. W., & Lemieux, T. (2008). *Regression discontinuity designs: a guide to practice*. Journal of Econometrics, 142.