基于 Harmonized CHARLS 数据的收入和健康恩格尔曲线

冯钰琪；Patrick McCarthy

（湖南大学经济管理研究中心，湖南省长沙市，410006）

摘要：中老年人群的健康问题受到越来越多的关注，越来越多的学者开始对中老年人群的健康消费进行研究。在本研究中，我们利用 Harmonized CHARLS 数据探讨收入与卫生支出之间的关系，发现收入对被调查者的健康支出有显著的影响，健康支出的收入弹性约为 0.023，这意味着医疗保健对被调查者来说仍然是一种必需品。另外，由于有相当一部分医疗支出为零，因此我们考虑使用 Tobit 模型进行回归分析受访者医疗支出的影响因素。研究结果表明，家庭收入对医疗支出的影响是显著的，养老保险收入和转移支付收入对医疗支出的影响也是显著的。

关键词：健康支出；收入；Tobit 模型；收入弹性

中图分类号： F 文献标识码：A

1. Introduction

[1]At present, China has become the country with the largest number of the older in the world and one of the countries with the fastest population aging. Although compared with many countries and regions in the world, such as Japan and Europe, China's aging problem has not appeared for a long time, but in recent years it has become more and more intense. [2]On the one hand, the average life expectancy in China has continued to increase. At the same time, due to the relevant policies implemented by China, the growth rate of the young population has been lower than the natural growth rate. Under the combined influence of these two factors, China’s aging problem has become more severe, it has attracted worldwide attention. Many people call it the “silver wave” or “population tsunami”.

[3]By around 2022, it was predicted that China's population over the age of 65 will account for 14% of the total population, realizing the transition to an aging society. From a long-term perspective, the older population will maintain a relatively high level both in terms of the overall number and the proportion of the population. The older are the most vulnerable and most in need of care among the entire population, so health and medical services are extremely important to them. It is worth noting that the older are under increasing pressure on health expenditure. [4]According to statistics from the Ministry of Health, the prevalence of chronic diseases in the older over 60 is 3.2 times that of the entire population, and the disability rate is 3.6 times that of the entire population. Moreover, the health resources consumed by the older are 1.9 times that of the average health resources consumed by the entire population.

Kanika Kapur et al. (2004) have examined that education, income, and wealth all affected medical care expenditures, although the effects of these variables differed across expenditure

[1] The data comes from the United Nations report.

[2] Derived from the "Statistical Bulletin on the Development of Health Services of our country" issued by the National Health Commission in recent years.

[3] Data from 2020 June 11, China Development Research Foundation released a report, "China Development Report 2020: China's aging population trends and policies."

[4] Data from the Ministry of Health.

categories[1]. Blendon (2002) has shown that the higher per capita income of household, the stronger the ability to purchase and pay for medical services, and the more demand there is[2]. For that reason, the health consumption of middle-aged and older people plays an important part of overall health consumption in China. Research on the influencing factors of health consumption of middle-aged and older people in China, especially the effect of income on health consumption, has become a critical topic for scholars.

In the existing research, there are a few studies on a particular group, especially on the middle-aged and older groups. However, it has been reported that the health expenditure of middle-aged and older people plays an important part of overall health expenditure in China. For that purpose, this paper focuses on vulnerable group over 45 years old, aiming to filling a gap in the literature by studying the relationship between income and health expenditure of that group. In addition, the existing research results mainly focus on the health expenditure itself, but a few studies integrate the health-care expenditure into the residents' overall consumption structure. Based on this, this paper will make a comparative analysis of food, transportation, alcohol and tobacco, communication, education, entertainment, clothing, fuels and other expenditures.

2. Literature Review

2.1 Studies of functional forms for estimating Engel Curves

Engel (1857) was first to propose that the proportion of food expenditure in the total expenditure declines as income increases, the law was later defined as "Engel laws"[3]. Engel curves reflect such a relationship, how the household expenditure on a particular good or service depends on household income. Engel curves can be regarded as a method to explore the relationship between the equilibrium quantity of a commodity purchased and the level of consumer income. Because of its simple form, Engel curves have been widely used in solving problems of consumer behavior.

A number of functional forms have been suggested and used as the basis for estimating Engel curves. In general, econometric studies of demand include both single equations and systems of demand equations. The single equation model is a commonly used method to estimate Engel curves. Working-Leser model, Heckman's sample selection model and the Tobit model are all single equation models. Working (1943) and Leser (1963) discussed the original form of the Working-Leser model firstly. In the Working-Leser model, each share of the food item is a linear function of the log of prices and of the total expenditure on all the food items under consideration. In many empirical works, zero consumption always exist. In order to correct zero consumption, Heckman's two-step model and the standard Tobit estimator are always applied. Heckman's two-step model has been proposed by Heckman (1976, 1979), who treats the selection problem as an omitted variable problem. The Heckman two-step model is based on the simultaneous estimation of two multiple regression models (an outcome equation and a selection equation) (Bushway et al., 2007; Barnighausen et al., 2011). This model has become the standard estimation procedure for empirical wage equations, has been criticized recently. The Tobit model is frequently used to analyze censored variables, which is subject to a known upper or lower bound. If zero consumption is assumed to be due to sample selection, Heckman's two-step is the appropriate model. The Tobit model simply captures the corner solutions for utility maximization.

2.2 The relationship between income and health expenditure

The research on health expenditure is relatively mature, many empirical analyses have confirmed that income is an important factor affecting health expenditure.

Baltagi et al (2017) used panel data on 167 countries from World Bank data set to investigate the relationship between healthcare expenditure and income[8]. After analyzing such many countries data, they provided an overall conclusion that health care is a necessity rather than a luxury at the global level. According to United Nations classification, China is classified as Asia-Pacific geo-political regional group. Different countries exhibit greatly different. For Asia-Pacific, the MG (Mean Groups) estimator shows an estimate of the income elasticity that is roughly 0.515, while when adding the public health expenditure rate, this estimate decreases to 0.505. The results support the hypothesis that for these countries health care is a necessity as opposite to a luxury. For WEOG (Western European and other group), the estimated income elasticity is roughly 0.691, which is also much lower than one. However, African and Latin American countries display a much larger income elasticity, equal to one when using the MG estimator. Although these countries with a larger coefficient than in wealthier countries, health care is also a necessity. The size of income elasticity depends on the position of different countries in the global income distribution, richer countries seem to show lower elasticity.

In China, researchers focus on a cross section of provinces in the early, with the deepening of research, they set about using panel data for research. Gu W. B. and Zhang D. G. (2008) used national data from 1985 to 2005 and co-integration and error Correction Model, analyzing the relationship between income of urban and rural residents and health expenditures in China[9]. In their paper, per capita disposable income is used to measure the income level of urban households, while the per capita net income is used to reflect the income level of rural households. According to the results, the per capita disposable income of urban households varies by 1%, and their annual health-care expenditure varies by 1.6847%. Compared with urban residents, there has slightly less impact of income on rural residents' health expenditure than urban residents, with the impact coefficient is 1.4688. The findings suggest that income growth will stimulate urban and rural residents to increase health care expenditure to some extent, and there is a long-term equilibrium relationship between rural residents' income and health care expenditure.

Considering the limitations of various commodity price data and the simultaneous endogeneity of the consumption share of various goods in the AIDS model, Zhou X. B. and Tian F. P. (2009) provided a nonparametric local linear estimation to the non-uniform panel data model[10]. Using the expenditure data and total expenditure data of eight categories of goods of urban residents in 9 towns in the Pearl River Delta region from 1993 to 2004 and 16 towns in the Yangtze River Delta region from 1994 to 2004, they established the Working-Leser model for non-uniform panel data of the two regions. Due to the year limit of data availability, the data they collected was non-uniform panel data. The findings suggest that in recent years, although the consumption share of health care products does not increase with the increase in total expenditure, it has become a necessity for people in the Pearl River Delta region. In their paper, they also show that the demand for goods in the categories of housing, health care, household equipment and services, clothing and entertainment and culture is elastic with a price elasticity of less than -1, which implies that when the price changes, the change in quantity demanded is greater than the change in price.

Tan T. , Zhang Y. Y. , et al. (2014) used the cross-sectional data of fixed observation points in the country in 2010 and the QUAIDS model to explore the influencing factors of rural household health expenditure and its demand elasticity[11]. The study confirmed that both family size and family income have a positive impact on health expenditure; the price elasticity of health

expenditure is -0.9522, thus health services still belong to life necessities on the whole, but as income increases, rural households will increase their consumption of health services. They also stratified the income of rural residents and calculated the non-compensation price elasticity of low-income rural households, middle-income rural households and high-income rural households respectively. The findings suggest that the absolute value of price elasticity of health expenditure in middle- and high-income rural households is higher than that of rural households with low income.

3. Research Hypothesis

This paper defines the utility function as:

$$u = lnU = \sum_{i=1}^{n} \beta_i lnq_i \tag{3.1}$$

Where u is the index of utility, q_i is the quantity of good i, and $0 < \beta_i < 1$.

The above equation is a version of the function that amounts to a logarithmic (hence monotonically increasing) transformation on term utility U. Thus, we obtain the first order condition

$$\frac{\partial u}{\partial q_i} = \frac{\beta_i}{q_i} > 0 \tag{3.2}$$

And it is larger than zero because of $\beta_i > 0$ and $q_i > 0$. Therefore, the utility function is a monotonically increasing function. That implies that the marginal utility of good i (MU_i) is positive. In addition, the function satisfies the law of diminishing marginal utility, i.e.

$$\frac{\partial^2 u}{\partial q_i^2} = -\frac{\beta_i}{q_i^2} < 0 \tag{3.3}$$

It is smaller than zero. As a result, the utility function is strictly quasi-concave.

The underlying assumptions of this utility function are homotheticity and additivity. The function is homothetic because utility rises by the scalar or a scalar proportion if each commodity is multiplied by scalar θ, i.e.

$$u + ln\theta \sum_i \beta_i = \sum_i \beta_i \ln(\theta q_i) \tag{3.4}$$

Also, it is strongly additive because the cross-partial derivative of the utility function is zero, i.e.

$$\frac{\partial^2 u}{\partial q_i \partial q_j} = 0 \tag{3.5}$$

Given the consumer's budget constraint,

(3.6)

$$m = \sum_{i=1}^{n} p_i q_i$$

Therefore, the augmented objective function is

$$L(q;\lambda) = \sum_i \beta_i ln q_i + \lambda(m - \sum_i p_i q_i) \tag{3.7}$$

where m is income, p_i is the price of good i, and λ is the Lagrange multiplier (or the marginal utility of income).

The first order conditions are

$$\frac{\partial L}{\partial q_i} = \frac{\beta_i}{q_i} - \lambda p_i = 0 \tag{3.8}$$

$$\frac{\partial L}{\partial \lambda} = m - \sum_i p_i q_i = 0 \tag{3.9}$$

According to the first order conditions, we find that the demand function is

$$q_i = \frac{\beta_i}{\sum_j \beta_j} \frac{m}{p_i} \qquad (i \in j) \tag{3.10}$$

Note that the demand function is homogeneous of degree zero in prices and income, by which we mean that if all the prices in MU equations and income increase in the same proportion, the quantity demanded is unaffected. This feature implies that a typical consumer is not fooled by his higher nominal income if prices increase in the same proportion.

Differentiating the demand function with respect to income and prices, we obtain cross-price elasticity:

$$\varepsilon_{ij} = \frac{\partial q_i}{\partial p_j} \frac{p_j}{q_i} = 0 \tag{3.11}$$

The cross-price elasticities are zero, implying that goods *i* and *j* are *grossly* independent. Note that the gross concept is based on the cross-price elasticity being equivalent to the total effect in the Slutsky equation.

Multiplying through the demand functions by p_i yields the Engel expenditure function

$$E_i = p_i q_i = \frac{\beta_i}{\sum_j \beta_j} m \tag{3.12}$$

The Engel expenditure function expenditure $p_i q_i$ to income m. This equation is linear in income and is sloped upward.

Based on the above analysis, two hypotheses to be tested are proposed and we will test them later:

Hypothesis 1: Holding all else constant, the demand for health services will decrease as their prices rise;

Hypothesis 2: Holding all else constant, respondents with higher income have higher health expenditure.

4. Model Setting

4.1 Model setting of Engel curves

In this study, we utilize Engel curves to estimate the relationship between income and health expenditure of respondents. Moreover, in order to examine the effect of income on health expenditure, we set the mathematical model and the model is expressed as follows.

$$\ln(Y_{it}) = \alpha_0 + \alpha_1 \cdot \ln(income_{it}) + X_{it} \cdot \theta + \varphi_t + \varphi_i + \epsilon_{it} \quad (4.1)$$

The subscript *i* represents different individuals, *t* represents the year; the dependent variable *Y* is the individual health expenditure and the number of doctor visits; income is the individual per capita household income; *X* is the individual characteristic variable; φ_t is the year fixed effect; φ_i is the individual effect. In order to eliminate the influence of heteroscedasticity, this paper takes the logarithm of health expenditure and income.

4.2 Model setting of Tobit model

4.2.1 Construction of Tobit regression equation

For censored data, when the left restricted point is 0 and there is no right restricted point, this model is the so-called "specification censored regression model", also known as Tobit model (Tobin, 1958). The Tobit regression model is commonly used in econometric study to model censored variables. Tobin observed that most households typically spend zero on major household durables such as cars or refrigerators over a given period of time. Thus, there are a lot of zero observations, which destroys the linear hypothesis. Therefore, general linear regression based on least square method cannot be used to analyze such data well.

Therefore, Tobin adopted a regression model that could cover non-negative expenditure when analyzing household durables consumption level and influencing factors, that is, Tobit model combining Probit model and multiple linear regression model.

The model settings are as follows.

$$y_i^* = X_i'\beta + u_i \quad (4.2)$$

$$u_i \sim N(0, \sigma^2) \quad (4.3)$$

$$y_i = \begin{cases} y_i^* & \text{if } y_i^* > 0 \\ 0 & \text{if } y_i^* \le 0 \end{cases} \quad (4.4)$$

When the latent variable y^* is less than or equal to 0, the explained variable y is equal to 0; When y^* is greater than 0, the explained variable y is equal to y^* itself, and it is assumed that the perturbation term u_i obeies a normal distribution with a mean of 0 and a variance of σ^2.

4.2.2 Tobit model estimation

Since OLS was used to perform linear regression on the whole sample, its nonlinear disturbance term would be included in the disturbance term, resulting in inconsistent estimation. Tobit proposed to use MLE to estimate the model.

First, the probability density function of the mixed distribution is derived, and then its logarithmic likelihood function can be expressed.

When y_i = 0,

$$\begin{aligned} P(y_i = 0 \mid X_i) &= P(y_i^* < 0 | X_i) = P(u_i < -X_i'\beta | X_i) \\ &= P(u_i / \sigma < -X_i'\beta / \sigma | X_i) \\ &= \Phi(-X_i'\beta/\sigma) \end{aligned} \quad (4.5)$$

When $y_i > 0$,

$$P(y_i > 0 \mid X_i) = P(y_i^* > 0|X_i) = 1 - P(y_i^* \le 0|X_i)$$
$$= 1 - P(u_i \le -X_i'\beta|X_i) = 1 - P(u_i / \sigma \le -X_i'\beta / \sigma|X_i)$$
$$= 1 - \Phi(-X_i'\beta / \sigma) = \Phi(X_i'\beta / \sigma) \quad (4.6)$$

The Probability Density function is

$$f(y_i|X_i) = \left[\Phi\left(-\frac{X_i'\beta}{\sigma}\right)\right]^{I_{y_i=0}} \left[\frac{1}{\sigma}\Phi\left(\frac{y_i - X_i'\beta}{\sigma}\right)\right]^{I_{y_i=0}} \quad (4.7)$$

Where, I is an indicator function，The value is 1 if the condition represented by the subscript is correct, and 0 otherwise.

The logarithmic likelihood function for the entire sample is as follows.

$$\log L = \sum_{i=1}^{n} \left\{ I_{y_i=0} \, ln\left[\Phi\left(-\frac{X_i'\beta}{\sigma}\right)\right] + I_{y_i>0} \ln\left[\frac{1}{\sigma}\Phi\left(\frac{y_i - X_i'\beta}{\sigma}\right)\right] \right\} \quad (4.8)$$

4.2.3 The hypothesis test of Tobit model

The hypothesis Test of Tobit model is realized through Likelihood Ratio Test (LR), and the original hypothesis of this Test is:

$$H_0: \beta = \beta_0 \quad (4.9)$$

The LR statistic is:

$$LR = -2(lnL_r - lnL_u) \sim \chi^2(j) \quad (4.10)$$

Where, lnL_r is the likelihood function value estimated by ML with constraints, and lnL_u is the likelihood function value obtained by ML without constraints. If H_0 is correct, $lnL_r - lnL_u$ should not be very large.

5. Results: presentation and discussion

5.1 Benchmark regression results

In this paper, we start the result discussion by examining the relationship between income and health expenditure. In the regression of panel data, we use pooled OLS and fixed effect model to estimate.

5.1.1 Using pooled OLS to estimate the Engel curve for health care demand

Firstly, we use pooled OLS to estimate the Engel curves for health care demand, and the result is reported in Table 5.1. In Table 5.1, it has been shown that there is no significance of the effect of per capita household income and household income on health expenditure although we control some variables.

Table 5.1 Estimation results for respondent's health expenditure

	Dependent Variable=log (respondent's health expenditure)			
	(1)	(2)	(3)	(4)
Pension income	0.024***			
	(0.007)			
Household		0.014		

income				
		(0.009)		
Income			0.014	
			(0.010)	
Transfer income				0.027***
				(0.006)
Control var.	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates.
*** p<0.01, ** p<0.05, * p<0.1
Data source: Harmonized CHARLS data

Then, we also use pooled OLS to explore estimation results for health expenditure of respondent's spouse, and the result is reported in Table 5.2. In addition, we explore the results for the number of doctor visits of respondent by using pooled OLS, and the result is reported in Table 5.3.

Table 5.2 Estimation results for health expenditure of respondent's spouse by using pooled OLS

Dependent Variable=log (health expenditure of respondent's spouse)				
	(1)	(2)	(3)	(4)
Pension income	0.031***			
	(0.007)			
Household income		0.036***		
		(0.009)		
Income			0.041***	
			(0.009)	
Transfer income				0.056***
				(0.006)
Control var.	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates.
*** p<0.01, ** p<0.05, * p<0.1
Data source: Harmonized CHARLS data

Table 5.3 Estimation results for the number of doctor visits of respondent by using pooled OLS

Dependent Variable= the number of doctor visits of respondent				
	(1)	(2)	(3)	(4)
Pension income	-0.003			
	(0.003)			
Household income		-0.003		
		(0.005)		
Income			-0.004	
			(0.005)	

Transfer income				0.011***
				(0.003)
Control var.	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates.
*** p<0.01, ** p<0.05, * p<0.1
Data source: Harmonized CHARLS data

5.1.2 Using fixed effect model to estimate the Engel curve for health care demand.

Then, we derive the relationship between respondent's income and health expenditure by using fixed effect model, the result is shown in Table 5.4. In Table 5.4, the dependent variable represents health expenditure, which has been taken logarithm. It has been shown that the coefficient of income is positive, and the effect of income on health expenditure is significant at 10% level after controlling the variables: age, married, work, household size, smoke, drink, overweight and recall score. Furthermore, it suggests that every 1% increase in income will increase health expenditure by approximately 0.023 %.

As for controlled variables, we derive that married and work have negative effect on health expenditure. Specifically, the respondent who had work experience in the past year spent approximately 45.3% less on health expenditure than without work experience. It has been reported that the respondent who has spouse spent about 31.9% less on health expenditure than without spouse. Also, as the respondent's age is increased by 1 year, health expenditure will increase by about 10.4%. We also derive that other controlled variables: household size, whether smoke before or not, whether drink before or not, whether over weighted or not and recall score have no significant effect on health expenditure.

Table 5.4 Estimation results for respondent's health expenditure by using fixed effect model

Dependent Variable= log (respondent's health expenditure)				
Variables	(1)	(2)	(3)	(4)
Pension income	0.026***			
	(0.009)			
Household income		0.020*		
		(0.011)		
Income			0.023*	
			(0.012)	
Transfer income				0.013
				(0.008)
Age	0.093***	0.104***	0.104***	0.104***
	(0.011)	(0.011)	(0.011)	(0.011)
Married	-0.311*	-0.322*	-0.319*	-0.327*
	(0.172)	(0.172)	(0.172)	(0.172)
Work	-0.441***	-0.453***	-0.453***	-0.448***
	(0.080)	(0.080)	(0.080)	(0.080)
Constant	Yes	Yes	Yes	Yes

Dependent Variable= log (respondent's health expenditure)				
Variables	(1)	(2)	(3)	(4)
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates. In addition to the control variables reported in this table, control variables reported in column (1), (2), (3) and (4) include household size, smoking, drinking, overweight and recall summary score.

*** p<0.01, ** p<0.05, * p<0.1

Data source: Harmonized CHARLS data

Furthermore, it has been shown that income elasticity of health expenditure is approximately 0.023, which implies that health care is still a necessity for respondents. Based on the assumptions we proposed before, we get such a hypothesis that income elasticity is equal to 1. However, it has been shown that income elasticity of health expenditure is approximately 0.023, which is not consistent with the above hypothesis. According to our tests, the income elasticity is not satisfied with hypothesis we proposed, however, health care is still a necessity for respondents. Although the effect of income on health expenditure is significant, health expenditure is not sensitive to changes in income in this study.

We also derive the relationship between income and the health expenditure of respondent's spouse by using fixed effect model, the result is shown in Table 5.5. In this table, it has been shown that respondents' income, whether it is pension income, household income, transfer income or per capita household income, has a significant effect on spouse's health expenditure, which implies that respondent with higher income might have stronger health awareness and have medical treatment promptly. In addition, household size has significant effect on the number of visiting doctor of respondents at 1% level. Specifically, the more family members, the lower the health expenditure of the respondent's spouse. Also, there is a significant effect of respondent's BMI (body mass index) on spouse's health expenditure. We speculate that this might be closely related to the respondent's family lifestyle.

Table 5.5 Estimation results for health expenditure of respondent's spouse by using fixed effect model

	Dependent Variable= log (health expenditure of respondent's spouse)			
Variables	(1)	(2)	(3)	(4)
Pension income	0.020**			
	(0.009)			
Household income		0.026**		
		(0.011)		
Income			0.032***	
			(0.011)	
Transfer income				0.051***
				(0.008)
Age	0.104***	0.113***	0.112***	0.112***
	(0.011)	(0.011)	(0.011)	(0.011)
Married	-4.473***	-4.484***	-4.480***	-4.514***
	(0.188)	(0.188)	(0.188)	(0.188)
Household size	-0.050***	-0.057***	-0.051***	-0.054***
	(0.019)	(0.019)	(0.019)	(0.019)
Overw	0.195**	0.202**	0.202**	0.206**
	(0.089)	(0.089)	(0.089)	(0.089)
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates. In addition to the control variables reported in this table, control variables reported in column (1), (2), (3) and (4) include work, smoking, drinking and recall summary score.

*** p<0.01, ** p<0.05, * p<0.1

Data source: Harmonized CHARLS data

Moreover, we also derive the relationship between income and the number of doctor visits during last month by using fixed effect model, the result is shown in Table 5.6. It has been shown that both respondent's per capita household income and household income have significant effect on the number of doctor visits of respondent at 10% significance level.

Table 5.6 Estimation results for the number of doctor visits of respondent by using fixed effect model

	Dependent Variable= the number of doctor visits of respondent			
	(1)	(2)	(3)	(4)
Pension income	-0.001			
	(0.004)			
Household income		0.012*		
		(0.006)		
Income			0.012*	
			(0.007)	
Transfer income				0.002
				(0.004)
Work	-0.120***	-0.124***	-0.123***	-0.120***
	(0.040)	(0.041)	(0.041)	(0.040)
Household size	0.018*	0.015	0.018*	0.018*
	(0.010)	(0.010)	(0.010)	(0.010)
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates. In addition to the control variables reported in this table, control variables reported in column (1), (2), (3) and (4) include age, married, smoking, drinking, overweight and recall summary score.

*** p<0.01, ** p<0.05, * p<0.1

Data source: Harmonized CHARLS data

Based on above analysis, it has been shown that hypothesis that when other conditions remain unchanged, respondents with higher income have higher health expenditure has been satisfied. In addition, it has been shown that income elasticity of health expenditure is approximately 0.023, which implies that health service is a normal good for respondents, not an inferior. However, it does not satisfy the hypothesis we have proposed that income elasticity of health expenditure is equal to 1.

5.2 Tobit analysis

The Tobit regression model is commonly used in econometric study to model censored variables. After analysis, we find that more than 75% of the variables of health expenditure are 0. At this point, we consider using Tobit model for regression. Firstly, we explore Tobit analysis of OLS model and report the results. Then, we conduct random Tobit analysis and marginal effects.

5.2.1 Tobit analysis of OLS model

Firstly, this paper uses Tobit analysis of OLS model to analyze the influencing factors of respondent' health expenditure. Tobit regression equation is shown in the following formula.

$$Y_{it} = \begin{cases} \alpha_{it} + \beta X_{jt} + \varepsilon & if\ \alpha_{it} + \beta X_{jt} + \varepsilon > 0 \\ 0 & if\ \alpha_{it} + \beta X_{jt} + \varepsilon \le 0 \end{cases} \quad (5.1)$$

Among them，X_{jt} is independent variable, represents the influencing factors of respondent' health expenditure, Y_{it} is dependent variable, represents the health

expenditure of respondent i in year t. β is the coefficient of regression, ε is stochastic disturbance. In order to eliminate heteroscedasticity of the data without changing the nature and relationship of the data, we take the natural logarithm of the continuous variable, and the corresponding Tobit regression equation is as follows.

$$Y_{it} = \alpha + \beta_1 lnX_1 + \beta_2 lnX_2 + \cdots\cdots + \beta_n lnX_n + \varepsilon \ (n = 1, 2, \cdots\cdots 10) \qquad (5.2)$$

In this formula, $X_1, X_2, \cdots, X_n$ respectively represent the explanatory variables used in Table 5.7.

Table 5.7 Estimation results for Tobit analysis of OLS model

Dependent Variable= log (respondent's health expenditure)				
Variables	(1)	(2)	(3)	(4)
log (household income)	0.129*			
	(0.078)			
log (pension income)		0.216***		
		(0.051)		
log (transfer income)			0.194***	
			(0.052)	
log (income)				0.132
				(0.083)
Control var.	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates. Control variables reported in all columns include age, work, married, household size, smoking, drinking, overweight, recall summary score and ADL score.
*** p<0.01, ** p<0.05, * p<0.1
Data source: Harmonized CHARLS data

In Table 5.7, this paper uses Tobit analysis of OLS model to analyze the influencing factors of respondent' health expenditure. We use household income, pension income, transfer income and per capita household income respectively as our independent variables, respondent's health expenditure as our dependent variable. Control variables including age, work, married, household size, smoking, drinking, overweight, recall summary score and ADL score. In order to eliminate heteroscedasticity of the data without changing the nature and relationship of the data, we take the natural logarithm of the continuous variable, including age, household size, recall summary score and ADL score.

According to Table 5.7, the results reveal that whether we use household income, pension income or transfer income as our independent variables, the effect of those measures of income on health expenditure is significant. Also, the coefficients of those measures of income are positive. However, for per capita household income, the effect is not significant. To be specific, the effect of household income on health expenditure is significant at 10% level, the effect of pension income and transfer income on health expenditure are both significant at 1% level. It should be noted that the coefficients of those measures of income cannot be directly used as the marginal effect of the explained variable, but can be used as the marginal effect of the latent variable. In the following section, we will discuss the marginal effect.

5.2.2 Random Tobit analysis

Then, this paper will explore random Tobit analysis to analyze the influencing factors of respondent' health expenditure. Specific regression results are shown in Table 5.8.

In Table 5.8, this paper uses random Tobit analysis to analyze the influencing factors of respondent's health expenditure. Using household income, pension income, transfer income and per capita household income respectively as our independent variables, respondent's health expenditure as our dependent variable. Control variables including age, work, married, household size, smoking, drinking, overweight, recall summary score and ADL score. Similarly, in order to eliminate heteroscedasticity of the data without changing the nature and relationship of the data, we take the natural logarithm of the continuous variable, including age, household size, recall summary score and ADL score.

According to Table 5.8, it has been shown that whether we use household income, pension income, transfer income or per capita household income as our independent variables, the effect of those measures of income on health expenditure is significant. Also, the coefficients of those measures of income are positive. Specifically, the effect of household income and per capita household income on health expenditure are both significant at 10% level, the effect of pension income and transfer income on health expenditure are both significant at 1% level. Similar to Tobit analysis of OLS model, it should be noted that the coefficients of those measures of income cannot be directly used as the marginal effect of the explained variable, but can be used as the marginal effect of the latent variable. In the following section, we will discuss the marginal effect.

Table 5.8 Estimation results for random Tobit analysis

Dependent Variable= log (respondent's health expenditure)				
Variables	(1)	(2)	(3)	(3)
log (household income)	0.146*			
	(0.078)			
log (pension income)		0.234***		
		(0.051)		
log (transfer income)			0.165***	
			(0.052)	
log (income)				0.151*
				(0.083)
Control var.	Yes	Yes	Yes	Yes
Constant	Yes	Yes	Yes	Yes
Observations	18,249	18,249	18,249	18,249

Notes: Standard errors are in parentheses below coefficient estimates. Control variables reported in all columns include age, work, married, household size, smoking, drinking, overweight, recall summary score and ADL score.
*** p<0.01, ** p<0.05, * p<0.1
Data source: Harmonized CHARLS data

6. Policy Implications

As far as income elasticity is concerned, the research conclusions of this article point out that for respondents, health service is still a necessity on the whole. Because of health expenditure is related to the survival and development of residents, it is necessary for the government to ensure

the quantity and quality of health services received by residents through corresponding policy support and financial support. For instance, it can consider increasing and improving the health facilities, and improving the quality of health services.

In terms of income level, this paper finds that the respondent's income level will have a significant positive impact on the health expenditure, and as income increases, respondents will increase their expenditure of health services. Therefore, the government should be committed to promoting the improvement of the income and living standards of the residents in both the rural and urban areas. Specific practices can be summarized as follows.

6.1 Improve the unified social security system for urban and rural areas

Firstly, improve the unified systems of basic medical insurance, serious disease insurance and basic old-age insurance for rural and non-working urban residents. Secondly, consolidate the direct settlement of medical insurance payments across the country. Establish and improve mechanisms for determining basic old-age insurance benefits for rural and non-working urban residents and making regular adjustments to basic old-age pensions. Do a better job of transferring and continuing social insurance relationships, and establish a public service platform for social insurance with the national government Affairs Service platform as the unified entrance. Build a multi-level old-age security system in rural areas and develop new models of diversified care services.

6.2 Improve the medical and health service system, especially in rural areas

Establish and improve relevant policies and systems, make community-level medical workers more attractive, and strengthen the training of medical and health professionals in rural areas. In addition, improve the conditions of township hospitals and village clinics, set up a sound system for the collection and transfer of medical waste in light of local conditions, improve our ability to prevent and treat chronic, occupational, endemic and major infectious diseases, strengthen mental health work, and advocate healthy birth and healthy rearing. Improve the operation mechanism of network services, encourage county hospitals and township hospitals to establish county medical community, and encourage large urban hospitals and county hospitals to establish peer assistance, mobile medical treatment and telemedicine mechanisms. Establish a tiered medical treatment system across the board and implement differentiated medical insurance payment policies. Also, establish and improve a nationwide fitness service system in light of local conditions.

7. Conclusions

As China is gradually entering an aging society, the health problems of middle-aged and older people have received more and more attention, and more and more scholars have begun to study the health consumption of middle-aged and older people.

In this study, we explore the relationship between income and health expenditure, finding that there has significant effect of income on health expenditure of respondents, and income elasticity of health expenditure is approximately 0.023, which implies that health care is still a necessity for respondents. Accounting for the difference for gender, educational level and hukou, we also conduct heterogeneity analysis for health expenditure. Moreover, the result implies that the effect of income on health expenditure is indeed different for different types of respondents. One possible reason is that respondents with higher income seem to be more health-conscious and more willing to spend for their own health than those with lower income.

In addition, because there are quite a few health expenditures that are zero, hence we consider using Tobit model for regression to analyze the influencing factors of respondent's health expenditure. We explore both Tobit analysis of OLS model and random Tobit analysis, and present the marginal effects. The results reveal that the effect of household income on health expenditure is significant, as well as pension income and transfer income. Based on the above analysis, we find that household income is an important factor influencing respondent's health expenditure, and health care is still a necessity for respondents. Therefore, it is necessary for the government to ensure the quantity and quality of health services received by residents through corresponding policy support and financial support. Based on this, we propose the following policy implications. Firstly, improve the unified social security system for urban and rural areas. Secondly, improve the medical and health service system, especially in rural areas. Thirdly, focus on alleviating structural employment problems.

8. References

[1] Kapur, K., Rogowski, J. A., Freedman, V. A., Adams, J. L., & Escarce, J. J. (2004). Socioeconomic status and medical care expenditures in Medicare managed care (No. w10757). National Bureau of Economic Research.

[2] Blendon, R. J., Schoen, C., DesRoches, C. M., Osborn, R., Scoles, K. L., & Zapert, K. (2002). Inequities in health care: a five-country survey. Health Affairs, 21(3), 182-191.

[3] Engel, E. (1857). "Die Produktions-und Consumtionsverhältnisse des Königreichs. Sachsen." Zeitschrift des Statistischen Büreaus des Königlich Sächischen Ministeriums des Innern, 8 and 9. (Reprinted in Engel (1895). Appendix I, pp. 1–54.)

[4] Banks, J., Blundell, R., & Lewbel, A. (1997). Quadratic Engel curves and consumer demand. Review of Economics and statistics, 79(4), 527-539.

[5] Newhouse, J. P. (1977). Medical-care expenditure: a cross-national survey. The Journal of Human Resources, 12(1), 115-125.

[6] Hitiris, T., & Posnett, J. (1992). The determinants and effects of health expenditure in developed countries. Journal of health economics, 11(2), 173-181.

[7] Moscone, F., & Tosetti, E. (2010). Health expenditure and income in the United States. Health economics, 19(12), 1385-1403.

[8] Baltagi, B. H., Lagravinese, R., Moscone, F., & Tosetti, E. (2017). Health care expenditure and income: a global perspective. Health economics, 26(7), 863-874.

[9] Gu, W. B., & Zhang, D. G. (2008). An Empirical Analysis of the Relationship between Urban and Rural Residents' Income and Medical Care Expenditure. Consumer economics (1), 43-46.

[10] Zhou, X. B., & Tian, F. P. (2009). Analysis of the difference in healthcare consumption between urban and rural residents in China: non-parametric estimation based on the panel data Engel curve model. Statistical Research (03),51-58.

[11] Tan, T., Zhang Y. Y, & He, J. (2014). Analysis on the Influencing Factors and Elasticity of Medical Consumption Expenditure of Chinese Rural Households. Journal of Shanghai University of Finance and Economics (03),63-69+112.

[12] Rous, J. J., & Hotchkiss, D. R. (2003). Estimation of the determinants. of household health care expenditures in Nepal with controls for endogenous illness and provider choice. Health Economics, 12(6), 431-451.

[13] Cleary, P. D., Mechanic, D., & Greenley, J. R. (1982). Sex differences in medical care utilization: an empirical investigation. Journal of Health and Social Behavior, 106-119.

[14] You, X., & Kobayashi, Y. (2011). Determinants of out-of-pocket health expenditure in China. Applied health economics and health policy, 9(1), 39-49.

[15] Wallace, R. B., & Herzog, A. R. (1995). Overview of the health measures in the Health and Retirement Study. Journal of Human Resources, S84-S107.

[16] Han, X., Qi, H. T. & Wang, X. H. (2019). Research on the impact of income on food consumption patterns of urban residents-based on two-stage EASI model estimation. Journal of Beijing University of Aeronautics and Astronautics (Social Sciences Edition) (02),92-98.

[17] Xue, Q. Z. (2009). An Empirical Study on Medical Consumption of Rural Residents in China (Master's Thesis, Nanjing Agricultural University).

[18] Li, Y. (2019). Research on China's household micro-consumption structure and characteristics (Ph.D. thesis, University of International Business and Economics).

Income and health care Engel curves in China: based on Harmonized CHARLS data

Yuqi Feng ; Patrick McCarthy

(Center for Economics, Finance and Management Studies, Hunan University, Changsha, 410006)

The health problems of middle-aged and older people have received more and more attention, and more and more scholars have begun to study the health consumption of middle-aged and older people. In this study, we explore the relationship between income and health expenditure by using Harmonized CHARLS data, finding that there has significant effect of income on health expenditure of respondents, and income elasticity of health expenditure is approximately 0.023, which implies that health care is still a necessity for respondents.

In addition, because there are quite a few health expenditures that are zero, hence we consider using Tobit model for regression to analyze the influencing factors of respondent's health expenditure. We explore both Tobit analysis of OLS model and random Tobit analysis, and present the marginal effects. The results reveal that the effect of household income on health expenditure is significant, as well as pension income and transfer income.

Keywords: Health expenditure; Income; Income elasticity; Tobit model; CHARLS