Does Education Level Influence the Consumption Behavior of Seniors in China?

Yijie Zhao

December 16, 2018

`

Abstract

This paper uses the data from the 2013 China Health and Retirement Longitudinal Study (CHARLS 2013)to test the impact of the education level of the seniors on their consumption level and living standards. The White heteroskedasticity consistent estimator is used to make the inference. With higher education, the seniors consumption level, living standard and the consumption of non-essential items have increased significantly, while the consumption of food and health-care remain relatively stable.

1 Introduction

The relationship between education and consumption is a popular topic being discussed in many economics researches. However, the consumption level and behavior of seniors are always being overlooked and stereotyped. Nowadays, the aging issue becomes more and more serious, especially in China. Seniors play an important role in society and economy in China. Moreover, the seniors who are currently in their 50-80s grew up during the turbulent times in China. The cultural revolution in the 1960s to 1970s had a devastating impact on the education system and limited the education and university enrollment. After the restoration of the education system in 1980, the number of university spots was still limited, and the competition for higher education was still extremely fierce. The education level of this group of seniors can lead to significant differences in their career and lifestyle. So we consider education level as an essential factor in their consumption behaviors. Since education can improve the value of human capital through knowledge and skills, and imperceptible influence consumer consciousness and behaviors, different education level can lead to diverse patterns of consumption. Therefore, this research will focus on analyzing how education level influences the current seniors' consumption in China.

This paper first gives a brief literature review about elder consumption and its possible relationship with education. Then we use the CHARLS(2013) to validate elder education's influence on senior consumption. Last, we run several tests to examine the possible heteroskedasticity and multicollinearity problems.

2 Literature Review

There are a number of competing claims in the theoretical and empirical literature about elders' consumption. For example, in medical decision making, older adults are more likely than younger adults to indicate that they would leave the medical decision up to the doctors instead of making them themselves(Steginga and Occhipinti (2002)). Another similar research reports that older adults, who perceive their time horizon as limited, place greater emphasis on emotionally meaningful goals (goals related to new information acquisitions). This age-related shift in goals could influence the type of information people use to make decisions

(affecting knowledge versus other types of information)(Fung and Carstensen (2003)).Despite growing interest in the topic of older consumers, there are still relatively few studies focus on the role the education play among the elders' consumption behavior.

Since the early 60s, when Schultz(Schultz (1963))theoretically stated education as an influencer among consumption, education has become a critical branch in consumption analysis. Empirically, for example, Li Yazhen et al. (2017)(Li and Zhang (2017))proves the influence of cognitive ability and mathematics ability on the high savings rate of the seniors, which explicitly indicates the relationship between the education and the consumption. Similarly, Ye Xiaomei et al. empirically shows that the level of education can significantly improve the elder individuals life quality, spiritually and culturally(Xiaomei and Wenyan (2016)) .

Although these papers show the elders consumption change in response to education level, they only focus on one particular consumption goods, possibly limited by the data. Therefore, this paper takes advantage of CHARLS as a complete and comprehensive database and aims to study the elder consumers behavior further.

3 Data

3.1 Data Description

China Health and Retirement Longitudinal Study 10 (CHARLS), led by Peking University's National Development Research Institute, is a two-year national household survey. This study surveys observations aging 45 or above, among overall 150 counties and 450 villages from 28 provinces. This paper selects the elder people over 50 and eliminates people over 100. This article matches two sub-database with observations ID and eliminates the outliers and missing values that are not informative. As a result, we exclude the sample information fail to be matched, ending up to 9618 after eliminating 8992 from 18610.

Table 1: Consumption in Categories

Variable	source	Mean	Std.Dev.	%
food	food	5246.89	6846.7	56.39
enjoy	servant, entertain, clothing, traveling, beauty	511.64	1812.42	5.50
daily	fuel , daily, heating, property, donations	588.97	4446.68	6.33
tech	communicate, utilities, transport, repair	1047.73	1315.18	11.26
train	training	355.33	1497.6	3.82
medical	medical, fitness	1554.17	5411.69	16.70

3.2 Dependent Variables

Consumption is measure in annual. We divided[1] detailed consumption into 6 categories, including : food consumption(food), enjoyment consumption(enjoyment), daily consumption(daily), high-tech and electronic products consumption(tech), education and training consumption(training), and medical care consumption(medical).

Table 1 summarizes some properties of the consumption among older people[2]. First, the average Engel coefficient from our sample is 56%, which means that seniors care about food consumption most. Second, the health-care consumption takes a large proportion since they face more disease health problems. In addition, the table shows that medical consumption has a high standard deviation, due to its low elasticity and significant variance. Third, in the last four categories, seniors consume relatively more in technology, indicating that the seniors have a greater acceptance of technology goods, which oppose our stereotype. The dependent variable shows in Table 2.

3.3 Independent Variables

Education is the independent variable. We use the highest level of education in CHARLES questionnaire which is captured by six dummy variables. Since older people rarely obtained education higher than the master in the data, we combine the sample with bachelor, master,

[1]How we divided shows in table 1.

[2]The last column is ratio each category consumption to overall consumption: $100\% * exp_i/exp$.

Table 2: Category Consumption in Percentage

Variable	Equation	Explanation
engal	food/exp*100%	Engel coefficient
enjoyper	enjoy/exp*100%	entertainment consumption %
dailyper	daily/exp*100%	daily consumption %
techper	tech/exp*100%	technology consumption %
trainper	train/exp*100%	training consumption %
medicalper	medical/exp*100%	health consumption %

Table 3: Description of Education

Dummy Variable	The Highest Level of Education	Number of People	%
0.*edu*	No formal education (illiterate)	2533	26.34
1.*edu*	Capable of reading and/or writing	1719	17.87
2.*edu*	Elementary school	2170	22.56
3.*edu*	Middle school	2032	21.13
4.*edu*	High school	970	10.09
5.*edu*	Bachelors degree /Masters degree/ Ph.D.	194	2.02

and PhD degree as one education status.Table 3 demonstrates the proportion of each education level in the whole sample.

We can see from the frequency distribution that the distribution of the population's education years is not uniform. Most seniors have lower education levels. However, through fitting analysis, the uneven distribution of samples rarely affect the results of the regression analysis. The regression results are still highly informative and representative.

3.4 Control Variables

Based on existing literature, we chose variables from the questionnaire that may affect personal consumption simultaneously, including gender(gender), hukou(hukou), communist

Table 4: Control Variables

Variable		Variable Description	
		0	1
Age	age	50-100	
Annual income per capital	lgincome	ln(income)	
Net assets per capita	lgwealth	ln(wealth)	
Gender	gender	female	male
Hukou	hukou	rural hukou	urban hukou
Communist Party	com	non-party member	party member
Marital status	mar	unmarried	married

party(com), marital status(mar), age(age), annual income per capita(income), and net assets per capita(income). We apply the logarithm to weaken the impact of singular values on regression results, which makes it easier to calculate and eliminate heteroscedasticity. Table 4 shows these control variables.

4 Empirical Setup and Result

4.1 Empirical Setup

For seniors, age 50 to 100 in China, their education background helped develop their consumption view. The level of consumption would be correspondingly increased with more education. For seniors with higher education level, they tend to consume more on non-essential products and services, such as entertainment, tourism, and training. In other words, a higher level of education will prompt consumers to increase consumption levels and optimize the consumption structure, and ultimately show an improvement in the quality of life and living standards.

We first investigate the effect of education level on total expense by estimating the following equation:

$$lgexp = \beta_0 + \beta_1 1.edu + \beta_2 2.edu + \beta 33.edu + \beta 44.edu + \beta 55.edu + \beta_i X_i + \varepsilon \quad (1)$$

The dependent variable $lgexp$ is the logarithm of annual consumption per capita, the independent variable.edu is a dummy variable denoting the education level of each individual, and X_i are individual controls including age, the logarithm of annual income per capita, the logarithm of net assets per capital and dummy variable for gender, hukou, communist party status, and marital status. ε is a disturbance term.

As older people with higher education level, they generally have higher reliance and requirements on the quantity and quality of goods and services. Among the subdivided consumption items, expenses on entertainment, travel, education, and training could reflect a high standard of living and consumption level. After considering the difference in annual income and net assets caused by different levels of education, we can infer that the direct impact of education on consumption can reflect the change of the individual's perception to consumption. For example, education could subtly and imperceptibly change the taste and style of a person's life. When purchasing goods, they will not only focus on the practicality of the goods but also consider some aesthetic aspects of goods. Higher education levels could also foster a more open-minded attitude to accept a broader range of goods and advanced high-tech products.

All in all, many researches have validated that the increase in income is positively related to the education level, as we did. It indicates education can influence consumption through income. Moreover, education can influence consumption by changing people's cultural tastes, values, and rationality.

We then examine the effect of education level on the weight of all subdivided consumption items respectively. In particular, we estimate the following equations:

$$engel = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \quad (2)$$

$$enjoyper = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \quad (3)$$

$$dailyper = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \quad (4)$$

$$techper = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \quad (5)$$

$$trainper = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \quad (6)$$

$$medicper = \beta_0 + \beta_1 1.edu + \beta_i X_i + \varepsilon \tag{7}$$

The dependent variables *engel* represents Engel coefficient, *enjoyper* represents the weight of entertainment consumption, *dailyper* represents the weight of daily consumption, *techper* represents the weight of technology consumption, *trainper* represents the weight of training consumption and *medicalper* represents the wealth of healthy related consumption. The independent variable *edu* is not a dummy variable here, it only shows the 6 steps of education level through 0,1,2,3,4,5, because we now care more about the effect of the hierarchy of education rather than the effect of a specific level of education respectively. X_i are individual controls including age, the logarithm of annual income per capital, the logarithm of net assets per capital and dummy variable for gender, hukou, communist party status and marital status. ε a disturbance term.

4.2 Heteroskedasticity and Multicollinearity test

After conducting the Breusch-Pagan test, we have to reject the hypothesis of homoscedasticity. Thus our model has the heteroskedasticity problem. Thus the usual OLS estimator will no longer remain efficient, and the usual inference procedures are no longer appropriate.

And as we do not know the exact variances of errors, we choose to use the White Heteroskedasticity Consistent Estimator. This estimator is based on the consistency of OLS estimator and allows us to approximate the squared residuals. In addition, as we have a large sample of 9618 observations, it is appropriate for us to use this method. Thus, after correcting the standard errors of our estimator, we can make a correct inference based on the results of the usual OLS estimators. Besides, we also check the variance inflation factor of the model to test for multicollinearity. From the test, we found that except for the educational dummy variables, most variables show little correlations. The indicators are satisfying, and multicollinearity is not a major concern here.

4.3 Empirical Result

Table 5(Appendix 1) reports the results on the effect of dummy education variables on total annual expense, which is consistent with our expectation. As the addition of control

variables, the effect of education on expense is weakened gradually.R-square and adjusted R-square are both increased from 5% to 16%. We recognize that the results of the fourth column is more reasonable and robustness, since too many control variables may cause the over-specified problem and bad control problem; and $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are all statistically significant at $p = 0.001$ level.

According to the result of the fourth column, the seniors who are capable of reading or/and writing has a marginal expense of 7% than illiterate seniors. Thus, the marginal expense of reading or/and writing competence level senior compares to no formal education level is 7.7%. Similarly, the marginal expense of elementary school level senior compares to no formal education level is 8.4%. The marginal expense of middle school level senior compares to no formal education level is 17.3%. And the marginal expense of high school level senior compares to no formal education level is 17.3%. The marginal expense of college education (Bachelors degree /Masters degree/Ph.D.) level senior compares to no formal education level is 46.8%. Besides, our table also shows that the annual income, net assets, hukou, and communist party status are all positively related with the total expense and the age and marital status are negatively related with the total expense. Most importantly, all coefficients, β_i of these control variables are statistically significant at $p = 0.001$ level. However,the β_i of gender is neither magnitude nor significant, which means the gender has little effect on the total expense.

The residual plot (Figure 1) and the normal QQ plot (Figure 2)(Appendix 2)of the model both look promising. Residuals do not show evident correlation with the fitted values, and the normality of the model seems convincing based on the plot.

Table 6(Appendix 3) reports the effect of education on the weights of subdivided consumption items. First of all, the effect of education on the weight of medical and health expenditure is neither the magnitude nor significant. However, its effect on Engel coefficient, entertainment consumption, daily expenditure, technology consumption and training consumption are all magnitude and significant. It is because that medical and health expenditure is unavoidable and inelastic, while the education level can greatly and significantly reduce the Engel coefficient and improve the seniors expenditure and consumption on goods and services in entertainment, travel, high-tech electronics, beauty products, property investment, cars and

so on. That is to say, seniors who have a higher level of education tend to have higher living standard and better consumption structure.

5 Conclusion and Discussion

5.1 Conclusion

With empirical evidence, this study reveals that the seniors consumption behavior corresponds to their former education. More importantly, individuals education level has a significant positive influence on consumption like entertainment, travel, high-tech electronics, beauty products, property investment, cars and so on, which aims to pursue higher living standard. This result provides the insight of long-term influence of education on consumption. More, many existing pieces of research overlooked the seniors consumption, which always being stereotyped. Thus, our research demonstrates that seniors also have diverse consumption views and behaviors influencing by early education.

5.2 Discussion

This paper demonstrates a robust relationship between education level and senior consumption. However, there are still some limitations in our model. First, the income and wealth, as the control variables in the analysis, can be correlated with each other, which might lead to multicollinearity problem. Second, this model may encounter the problem of endogeneity. Although we already have controlled lots of variables in this study, there are still other variables related to education and influenced expenses, such as the income of parents and number of children. Since we cannot control all related variables, the omitted variable bias may occur. Moreover, CHARLS collected the data by surveys, which can raise the attenuation bias that interviewees may provide dishonest answers. Third, because of the political turmoil in the 1960s, the education level of current seniors in China was affected. The average education level of them is relatively low and the amount of people who received high education is smaller. Thus, the small sample with the education above high school can cause the systematic error in the regression result, which is irreducible.

References

Fung, H. H. and L. L. Carstensen (2003). Sending Memorable Messages to the Old: Age Differences in Preferences and Memory for Advertisements. *Journal of Personality and Social Psychology 85*(1), 163–178.

Li, Y. and C. Zhang (2017). Cognitive ability and consumption: a new perspective to understand the high savings rate of the elder people. *Economics Dynamics 3*, 65–75.

Schultz, T. W. (1963). *The Economic Value of Education*. Ph. D. thesis.

Steginga, S. K. and S. Occhipinti (2002). Decision making about treatment of hypothetical prostate cancer: Is deferring a decision an expert-opinion heuristic? *Journal of Psychosocial Oncology 20*(3), 69–84.

Xiaomei, Y. and L. Wenyan (2016). Can education improve the quality of life of elder people? - Evidence based on panel data from CLHLS(2005-2011). *Educational Economic Review 4*, 34–45.

Appendix 1 Table 5 the Effect of Dummy Education Variables on Total Annual Expense

VARIABLES	(1) lgexp	(2) lgexp	(3) lgexp	(4) lgexp	(5) lgexp	(6) lgexp	(7) lgexp	(8) lgexp
1.edu	0.122***	0.094***	0.085***	0.077***	0.070***	0.075***	0.071***	0.070***
	(0.024)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)	(0.023)
2.edu	0.180***	0.115***	0.094***	0.084***	0.070***	0.075***	0.069***	0.069***
	(0.022)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.022)
3.edu	0.265***	0.150***	0.110***	0.089***	0.057**	0.062***	0.051**	0.051**
	(0.022)	(0.022)	(0.021)	(0.022)	(0.022)	(0.022)	(0.023)	(0.024)
4.edu	0.443***	0.253***	0.194***	0.173***	0.110***	0.115***	0.095***	0.095***
	(0.028)	(0.027)	(0.027)	(0.027)	(0.028)	(0.028)	(0.029)	(0.030)
5.edu	0.937***	0.617***	0.481***	0.468***	0.359***	0.365***	0.319***	0.319***
	(0.047)	(0.047)	(0.046)	(0.046)	(0.048)	(0.048)	(0.050)	(0.050)
lgincome		0.160***	0.127***	0.128***	0.117***	0.118***	0.117***	0.117***
		(0.008)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
lgwealth			0.110***	0.107***	0.103***	0.102***	0.102***	0.102***
			(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
age				-0.003***	-0.004***	-0.005***	-0.006***	-0.006***
				(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
hukou					0.154***	0.153***	0.148***	0.148***
					(0.020)	(0.020)	(0.020)	(0.020)
mar						-0.085***	-0.089***	-0.089***
						(0.027)	(0.027)	(0.027)
com							0.098***	0.098***
							(0.025)	(0.025)
gender								0.001
								(0.016)
Constant	8.645***	7.319***	6.469***	6.708***	6.911***	7.037***	7.084***	7.084***
	(0.015)	(0.065)	(0.076)	(0.097)	(0.102)	(0.108)	(0.109)	(0.109)
Observations	9,618	9,618	9,618	9,618	9,618	9,618	9,618	9,618
R-squared	0.050	0.121	0.150	0.151	0.156	0.157	0.159	0.159
adjusted R-squared	0.0497	0.120	0.150	0.151	0.155	0.156	0.158	0.158

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix 2

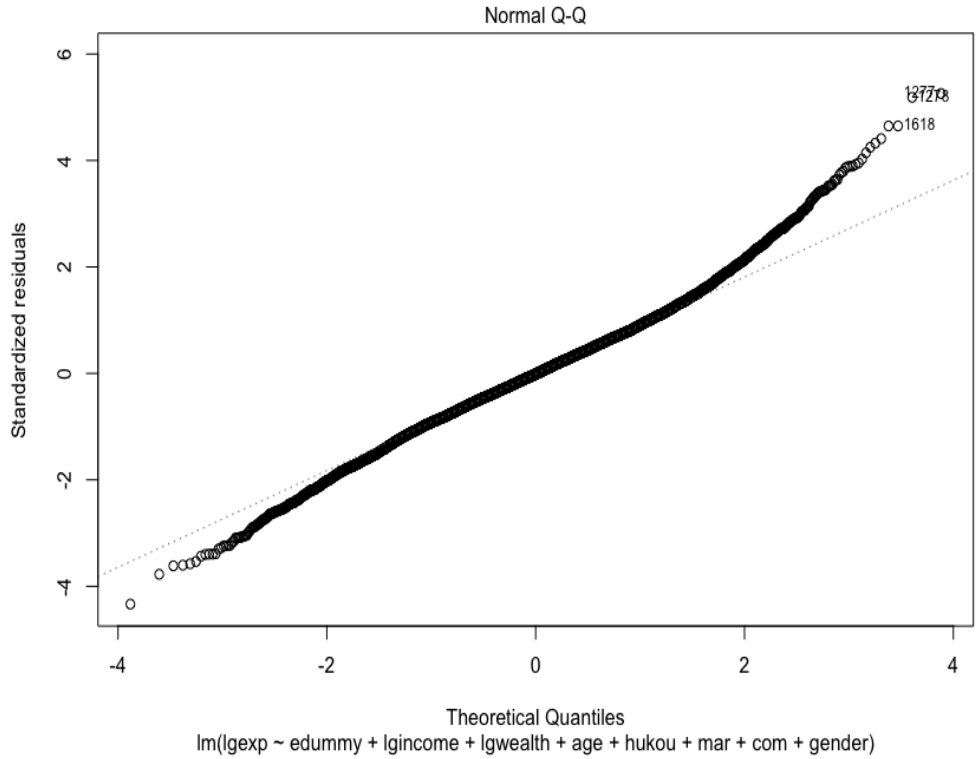

Figure 1. Residual Plot of the model

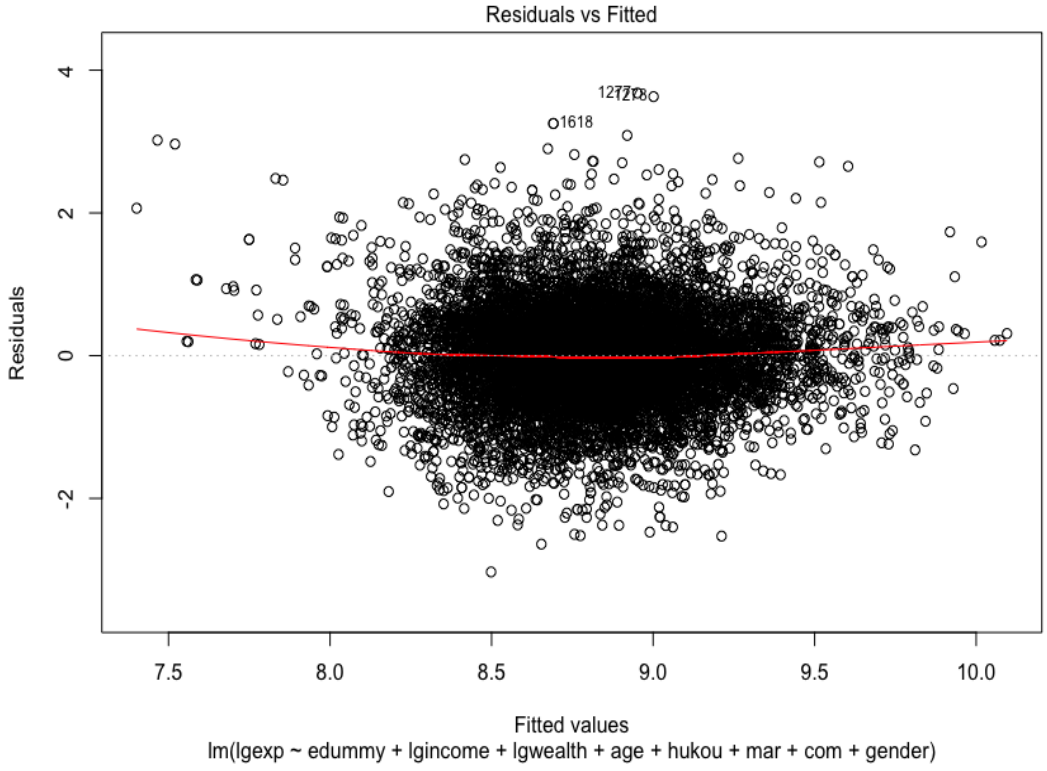

Figure 2. Normal QQ plot

Appendix 3

Table 6 Effect of Education on the Percentage of Subdivided Consumption

VARIABLES	(1) engel	(2) enjoyper	(3) dailyper	(4) techper	(5) trainper	(6) medicalper
edu	-1.145***	0.410***	0.185***	0.271***	0.341***	-0.062
	(0.203)	(0.067)	(0.071)	(0.096)	(0.095)	(0.160)
lgincome	-0.776***	0.564***	0.080	0.536***	-0.051	-0.354**
	(0.208)	(0.060)	(0.076)	(0.091)	(0.088)	(0.171)
lgwealth	-1.221***	0.761***	0.160**	0.856***	0.103	-0.659***
	(0.200)	(0.070)	(0.080)	(0.090)	(0.084)	(0.161)
age	-0.122***	-0.050***	0.059***	-0.088***	-0.146***	0.347***
	(0.029)	(0.010)	(0.011)	(0.013)	(0.014)	(0.024)
hukou	-4.009***	1.653***	2.339***	-0.662**	-0.172	0.851
	(0.653)	(0.243)	(0.239)	(0.305)	(0.313)	(0.533)
mar	-0.945	-0.276	-0.507	-0.253	-0.682**	2.663***
	(0.833)	(0.246)	(0.338)	(0.355)	(0.335)	(0.706)
com	-0.412	0.779***	0.005	0.053	-0.362	-0.064
	(0.757)	(0.281)	(0.270)	(0.357)	(0.338)	(0.613)
gender	1.961***	-0.422***	-0.702***	-0.132	-0.077	-0.628
	(0.495)	(0.158)	(0.172)	(0.234)	(0.231)	(0.398)
Constant	88.325***	-5.381***	0.272	5.144***	12.876***	-1.236
	(3.383)	(1.202)	(1.251)	(1.428)	(1.616)	(2.764)
Observations	9,618	9,618	9,618	9,618	9,618	9,618
R-squared	0.032	0.095	0.028	0.034	0.021	0.035
adjusted R-squared	0.0310	0.0943	0.0273	0.0329	0.0204	0.0342

Robust standard errors in parentheses

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

Appendix 4

```
library(foreign)
library(readxl)
library(dummies)
library(devtools)
library(broom)
library(readxl)
library(car)
library(het.test)

data <- read_excel("data.xlsx")
summary (data$edu)
data$edummy <- as.factor(data$edu)
data$edummy <- relevel(data$edummy, ref = '0')

lm1 <- lm(lgexp~ edummy, data = data)
summary (lm1)
lm2 <- lm(lgexp~ edummy + lgincome, data = data)
summary (lm2)
lm3 <- lm(lgexp~ edummy + lgincome + lgwealth, data = data)
summary (lm3)
lm4 <- lm(lgexp~ edummy + lgincome + lgwealth + age, data = data)
summary (lm4)
lm5 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou, data = data)
summary (lm5)
lm6 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar, data = data)
summary (lm6)
lm7 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar + com, data = data)
summary (lm7)
lm8 <- lm(lgexp~ edummy + lgincome + lgwealth + age + hukou + mar + com + gender, data
= data)
summary (lm8)


lm01 <- lm(engel~ edummy, data = data)
summary (lm01)
lm02 <- lm(engel~ edummy + lgincome, data = data)
summary (lm02)
lm03 <- lm(engel~ edummy + lgincome + lgwealth, data = data)
summary (lm03)
lm04 <- lm(engel~ edummy + lgincome + lgwealth + age, data = data)
summary (lm04)
lm05 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou, data = data)
summary (lm05)
```

```
lm06 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar, data = data)
summary (lm06)
lm07 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar + com, data = data)
summary (lm07)
lm08 <- lm(engel~ edummy + lgincome + lgwealth + age + hukou + mar + com + gender, data
= data)
summary (lm08)


par(mfrow=c(2,2))
plot(lm8)
par(mfrow=c(2,2))
plot(lm4)

vif(lm8)
bptest(lm8)
covmtx <- hccm(lm8)
coeftest(lm8, vcov = covmtx)

vif(lm4)
bptest(lm4)
covmtx <- hccm(lm4)
coeftest(lm4, vcov = covmtx)
```