ATLANTIS PRESS
Advances in Economics, Business and Management Research, volume 87
5th Annual International Conference on Management, Economics and Social Development (ICMESD 2019)

Does Health Insurance Help the Aged? Evidence from China

Zhu Li[a,*] and Yu-Xue Cheng[b]

China University of Petroleum Beijing, No. 18, Fuxue Road, Changping District, Beijing, 102249, P.R. China

[a]cruise_lee@outlook.com, [b]chengyuxue79@yeah.net

*Zhu LI

Keywords: Impact evaluation, Health assessment, Difference in difference, Propensity score match, Lasso regression.

Abstract. This paper discussed the effect of a health insurance program, New Cooperative Medical Scheme (NCMS) from China, on the elderly's reported health status. Using a Difference-in-Difference strategy, this paper finds evidence that the program has strengthened elders' daily activity livings and improved access to medical facility for those participants to secure on-time help. Issues regarding adverse selection is also discussed.

1. Introduction

Public programs are usually employed to help certain beneficiaries and reach specific societal goals. For a better understanding of whether those programs work satisfyingly, impact evaluation which often involves mixed methods will help policymakers' decision [1]. As a part of national public programs, health insurance plays a crucial role to improve the health condition of the poor and the old. For instance, to investigate the influence of health insurance in US, a literature specialized the effect on newly insured elders around age 65 showed the association between continuous insurance coverage and significantly fewer but a significant upward shift in the distribution of health states from good to poor among survival people [2].

There were also some developing country experiences in health insurance and related pieces of literature. Many of them focused on the effect on financial burdens like out-of-pocket expenditure and catastrophic expenditure [3,4]. The linkage between insured people and their characteristics also elicited intensive discussion. In a China-related study, using census data of health facility and surveys, the authors found some evidence to indicate lower enrollment rates among the poor households and higher enrollment rates among the households which contained chronically sick members. China's New Cooperative Medical Scheme (NCMS), which launched in 2003, had attracted many researchers' attention and investigation because of its unprecedented transformation on social resources.

The NCMS was designed to improve financial support for rural health services against serious illness. The enrollment for the program is optional, covering all rural county-level. Some study concerned the program based on the voluntary system would elicit underlying adverse selection. However, the high-level participation and relatively generous subsidies in NCMS seemed partly address the problem [5]. However, until now, there were still mixed results from empirical researches studied on the improvement of health status in the rural area with NCMS which start in 2003. From the China Health and Nutrition Survey (CHNS), although the authors employed multiple estimation strategies includes IV, PSM, fixed effect, PSM-DD to correct the potential selection bias, no evidence proved the policy improved health status measured by self - reported health status and by sickness or injury in the past four weeks [6]. Later, using the Chinese Longitudinal Healthy Longevity Survey(CLHLS), related works find insignificant health effect but accentuate the importance of susceptibility to health risk for the health effect on rural elderly population [7]. The complex results above should be paid more attention and need further discussion. Also, it was indicated more studies should focus on the real health effects of existed health scheme because the rural elderly health status was evidently worsened than urban's [8,9].

This study investigates the impact evaluation of the different outcomes of health conditions in the population who enrolled NCMS. For the evaluation analysis, a commonly accepted approach was employed in the literature, that is propensity score matching with difference-in-difference (PSMDD), which could remove partly selection bias. From the results by using difference-in-difference and propensity score matching with difference-in-difference methods, few significant treatment effects of NCMS were identified on various health outcomes of the elderly. Therefore, there is a finding which suggests selection concerns from unobservable covariates in the last discussion.

This paper consists of four parts. Section I introduce the insurance program. Next section discusses the data. Section III provides the empirical strategies and result analysis with alternative explanations. The final section contains a summary and discussion.

2. Data Sources and Variables

Table 1. Summary statistics

Variables	Pre-treatment year (wave 2005)			Post-treatment year (wave 2008)		
Full sample	Treated	Control	T-test	Treated	Control	T-test
N=2554	N=1672	N=882		N=1672	N=882	
Self-reported health (SRH)	0.52 (0.50)	0.51 (0.50)	0.56	0.42 (0.49)	0.45 (0.50)	0.19
Ability of daily live(ADL)	0.92 (0.27)	0.93 (0.25)	0.37	0.91 (0.29)	0.88 (0.32)	0.05 **
Comparative self-reported health (CSRH)	0.11 (0.31)	0.10 (0.30)	0.73	0.09 (0.29)	0.10 (0.30)	0.41
Age	79.85 (10.45)	80.44 (10.34)	0.17	83.00 (10.47)	83.59 (10.36)	0.17
Gender	0.45 (0.50)	0.48 (0.50)	0.23	0.45 (0.50)	0.48 (0.50)	0.19
Smoke	0.23 (0.42)	0.27 (0.44)	0.08 *	0.20 (0.40)	0.20 (0.40)	0.85
Drink	0.24 (0.43)	0.26 (0.44)	0.30	0.19 (0.39)	0.19 (0.40)	0.60
Exercise	0.28 (0.45)	0.26 (0.44)	0.35	0.24 (0.43)	0.29 (0.45)	0.01 ***
Schooling	1.63 (2.71)	1.79 (2.99)	0.20	1.63 (2.71)	1.79 (2.99)	0.20
Enough money	0.70 (0.46)	0.71 (0.45)	0.46	0.70 (0.46)	0.72 (0.45)	0.26
Revenue	7458.00 (20653.32)	10391.77 (24086.33)	0.00 ***	15629.39 (21757.29)	18817.91 (36083.31)	0.00 ***
Marriage	0.46 (0.50)	0.40 (0.49)	0.00 ***	0.40 (0.49)	0.37 (0.48)	0.23
Cure on time	0.86 (0.34)	0.85 (0.36)	0.32	0.92 (0.27)	0.83 (0.37)	0.00 ***
Medical cost	643.71 (1556.41)	841.6 (2134.07)	0.01 ***	1029.45 (2935.45)	1444.91 (4555.87)	0.01 ***
Sisters or brothers	3.31 (1.96)	3.19 (2.07)	0.17	3.27 (1.97)	3.07 (1.95)	0.01 ***

Notes: There are standard errors of different groups' participants in two periods. In the brackets two-sample t-test between treated and control groups. P < 0.01 ***; P < 0.0.5 **; P <0.1 *.

The data used in this study is from the Chinese Longitudinal Healthy Longevity Survey (CLHLS) in the websites [10,11]. The CLHLS is a national survey, launched in 1998, which paid attention to the representation of the oldest-old. The goal of this research is to get a better understanding of the various determinants of people's healthy longevity. The period of the data is from 1998 to 2014 with 88% average response rates. They collected data with practical procedures and constructed meaningfully comparable subsamples in the survey [12,13], with randomly selected female and males from counties. Therefore, it could be a qualified natural experiment.

Since the insured population rate in NCMS rose to 90% in 2011, the unbalanced ratio between insured elders and uninsured ones may confuse the comparability of these two groups in this research. Therefore, data on health outcomes based on 2005-2008 waves were used in this study. For sample selection, the involved respondents who died before 2008 were excluded in the estimation to ensure the sample numbers who have complete attributes in the study. The elderly who already insured before 2005 were excluded from the sample, and the samples in the rural area were kept. Finally, the dataset consisted of a balanced panel, including 2554 rural elders in waves of 2005 and 2008.

To analysis health outcomes based on 2005-2008 waves, the work used measures of health in respondents to assess whether salient health impact occurred in the elderly after they enrolled NCMS. Three dependent variables used as measurements on sample's health status were *self-reported health* (SRH),which has been proved by numerous studies to be a good measurement of health, and it is strongly associated with subsequent health outcomes such as mortality [15,16]; *activities of daily living* (ADL), which has been considerable practical value as a longitudinal measure, according to the ADL index of Katz [17]; *comparative self-reported health* (CSRH), An evaluation of the people's self-assessed health status relative to one' status last year concerning the subjectivity and age- sensitivity of self-reported health from different participants.

Several potential health indicators to describe people with demographic and socioeconomic features, which could possibly confound the health outcomes including *province, age, gender, education, smoke, drink, exercise, revenue, marriage, medical costs, the number of sisters or brothers* and other detailed questions about their life quality, which were denoted as personal features in the regression model. The groups of treated and control are divided based on the enrollment in 2008. There were 2554 people in this sample with 1672 elders enrolled NCMS as the treated group, and 882 continued uninsured as control one.

Table 1 displayed the descriptive statistics of the variables before and after the treatment year, which is the time the elderly enrolled in NCMS. Moreover, in 2005, the two-sample t-tests between groups indicated that there was no significant difference in health measurements initially. However, it is obvious to observe that there were significant differences in *smoke, revenue, marriage* and *medical cost* in 2005, which means married elders may associate with higher enrollment in NCMS, the low-revenue also correlated with whether elderly choose to enroll. People enrolled in NCMS had significantly lower medical cost than non-enrolled which could have been the possible motivation because of NCMS's help in financial support (During the period covered by the study (2005–2008), the minimum requirement for the household contribution was 20 RMB per person, and the subsidy payment was 40 RMB per person. In 2008 the standard subsidy level from central and local governments rose to 80 RMB per person.).

In 2008, for health outcomes, participants of NCMS maintained their ADL capability better than the uninsured. The participants started to take significantly less *exercise* than unenrolled elderly. The table also showed that the participants of NCMS received more *cure on time* than the control group, which give evidence in the moderate effect on health facilities. From 2005 to 2008, insured aged people have increased by 60% in average *medical cost* while the uninsured had increased by 70%. In both years, the uninsured spent much more on medical services than insured.

3. Empirical Framework

This paper aims to define the impact of insurance program on three health outcomes for the elderly. However, I should acknowledge the problems inside the models first. The selection effect: although the government regulate some rules to prevent adverse selection in NCMS (To reduce the risk of adverse selection and encourage the enrollment, the council state ruled that the financial support would offer to

the county only where the enrollment rate reached 80%. Because of the low premiums and active mobilization from local governments, some evidence indicated the adverse selection is existed but not significant [18].), people who care about their health status and future medical burden will incentive themselves participating in this voluntary system; data attrition could occur to cause sample selection bias, because I use the answers from alive people in both interviews while the mortality rate for three years is considerable.

3.1. Difference in Difference (DID) Framework

3.1.1. Identification

Through double difference (before and after the NCMS impact in participants), the treatment effect of the health insurance could be calculated by the two-stage data. The performance bias between the insured elders and uninsured ones could be corrected based on the assumption that unobserved heterogeneity is time invariant, which is called "parallel trend", through differencing. One of the advantages could help the evaluation from DID is that the convenience to add additional covariates in the framework [19]. Therefore, this study attempted to use DID to remove partly unobserved bias in the elderly

Considering some omitted variables in the study like province, family and individual heterogeneity that may influence the regression results with bias, the fixed effect model was employed to control the regional difference. The other covariates in this model were all the common characteristics in the other discussion about how the NMCS affect health outcomes. Notably, the model tested the validity of the effect from covariates involved which have the salient difference between participants and non-participants in 2005 as summarized above. After adding the province fixed effect and other observables to the DID framework, the generalized equation became Eq. 1:

$$Y_{iiiiii} = \alpha\alpha\alpha\alpha_{iiiiii} + \lambda\lambda_{ii} + \theta\theta_{ii} + \delta\delta_{ii} + \xi\xi\xi\xi + \varepsilon\varepsilon_{iiiiii} \quad (1)$$

Where Y_{iiiiii} is the health status outcomes of two groups' elderly in province p for three years. $\lambda\lambda_{ii}$ is the time fixed effect; $\theta\theta_{ii}$ is the group fixed effect; $\delta\delta_{ii}$ is the province fixed effect, based on the question on the survey; $\xi\xi$ is a range of observable covariates from the interviewing, $\varepsilon\varepsilon_{iiiiii}$ represents the error term.

Besides, the critical assumption to an unbiased outcome is that the treated and control groups were supposed to have the common trend in their health status, that is, participants and non-participants getting their health worse or better in the same level. Therefore, the model consisted additionally observable time-varying variables which may affect the health outcomes.

3.1.2. Data Results and Analysis

Table 2 reports the outcomes in DD estimates. The coefficients in *SRH* and *CSRH* were negative but insignificant while the *ADL* indicators showed a significantly upward around 0.03 in *ADL* scores for the participants. In the health access & service part, there were actively significant effect for the insured elders in the frequency and satisfactory of *cure on time*. Besides, the *medical cost* of the elderly would increase by 14% in significance if one was insured in 2005.

Panel A in column (1) is the outcome based on observable variables except the four initially different personal characteristics (smoke habit, revenue, the number of sister and brothers and marriage situation), which indicates the significantly positive effect of health insurance on ADL but few evidence of the association in good *self-reported health* and *comparative ones* with the NCMS. As for panel B in column (1), the improvement of cure speed is showed strongly caused by health insurance, the medical cost also was evident with a compelling P value - 0.05, as the primary goal of the program is. From column (2) to column (4), the difference of smoke habits between treated and control group made a moderate but positive role to explain the SRH as well as the medical cost. However, the revenue, which has salient different between groups did not help in the model. The relationship helps to explain the underlying association in medical cost. In column (5), the parameter got promoted in medical cost, which meant the model obtained more accuracy to understand the direct influence of NCMS like financial subsidy. In sum, from all columns above, the results of comparative self-reported health kept confound and may indicated no

association with insurance enrollment while that of the objective measurements provides strong evidence in association.

Two potential problems which could disrupt the ideal results are inadequate and inaccurate independences, and the existence of adverse selection. Firstly, the availability of data was limited and the health problem need more explanatory indicators. Besides, both of the statistical significance of differences and the improvement of the regression results when concerning additional four variables indicated the existence of heterogeneity between participants and non-participants of the NCMS in this study. Therefore, the distribution based on the enrollment could not be treated as randomly selected groups. The reason behind may be the adverse selection which consists of complicated motivations from the participants.

Table 2. DID estimation

	1	With smoke	With revenue	With relationship	All
	(1)	(2)	(3)	(4)	(5)
Panel A: Health outcomes					
SRH Self-reported health	-0.04 (0.03) [0.17]	-0.04 (0.03) [0.16]	-0.04 (0.03) [0.17]	-0.04 (0.03) [0.17]	-0.04 (0.03) [0.16]
ADL Activities of daily living	0.03** (0.02) [0.03]	0.03** (0.02) [0.03]	0.03** (0.02) [0.03]	0.03** (0.02) [0.03]	0.03** (0.02) [0.03]
CSRH Comparative Self-reported health	-0.01 (0.02) [0.41]	-0.01 (0.02) [0.41]	-0.01 (0.02) [0.41]	-0.01 (0.02) [0.43]	-0.01 (0.02) [0.43]
Panel B: Health access & service					
Log medical cost	0.14* (0.07) [0.05]	0.14** (0.07) [0.04]	0.14* (0.07) [0.05]	0.14** (0.07) [0.04]	0.15** (0.07) [0.03]
Cure on time	0.07*** (0.02) [0.00]	0.07*** (0.02) [0.00]	0.07*** (0.02) [0.00]	0.07*** (0.02) [0.00]	0.07*** (0.02) [0.00]
Time, group FE	Yes	Yes	Yes	Yes	Yes
Regional FE	Yes	Yes	Yes	Yes	Yes
Personal features	Yes	Yes	Yes	Yes	Yes
Smoke	No	Yes	No	No	Yes
Log revenue	No	No	Yes	No	Yes
Sister, brother & Marriage	No	No	No	Yes	Yes

Notes: All the value shown in the table is the NCMS effect on health-related outcomes. Standard errors reported in parenthesis. P-values are in square brackets, and the significance of 0.01,0.05,0.1are denoted by *, **, ***. Columns (2)(3)(4) were divided by the four initially different features into three group to regress respectively in. The regression with all variables written in the column (5). Personal features involved several characteristics which were not significantly different through two-sample t test from the statistical summary above.

In the study, self-selection could non-randomize the treatment I studied above because the statistical summary showed obvious difference which may happened in consistence with the selection effect. In sum, purposive program placement and self-selection, which is the primary concerns in this study, could cause the non-randomization for Treatment assignment. For example, the negative effect of NCMS on the self-reported health is counterintuitive, and the positive impact should be tested. Although the unobservable variables were impossible to recollect, the selection problem could be partly reduced by statistical method like sample matching.

Most self-selection in the real world is based on many individual characteristics. In the next minor section, a statistics method was employed, trying to redesign the experiment and eliminate some selection bias in impact evaluation.

3.2. Propensity Score Matching (PSM) and PSM-DID

Given potential issues concerning selection bias, another method was used in this section. For this task, the study presented a two-step procedure. First, the propensity score was estimated by using a logit model and rematch the samples in this experiment. Then a DID structure using the kernel propensity score as a weighting metric was presented.

To reach an unbiased average effect for NCMS enrollment, some inherent characteristics that could affect the participation are constructed in the model as conditional configuration [20]. For observational studies, matching could statistically help to make matched control group more similar with the participants. PSM condenses a set of covariates to a scalar to calculate every elderly's probability on the enrollment in the NCMS where the model is usually used to transfer various potential characteristics to the probability of participation.

One of the most critical assumptions in PSM is the *common support* because only in the area of common support can inferences be made about causality. All the characteristics were selected carefully in this study to satisfy the region of common support, the overlap area of the treated and control groups' p-score distribution. The extreme characteristics were dropped to prevent the match from the bias caused by weak common support [21]. The samples were re-matched the samples with their individual information like *age, gender and schooling*, the previous life habits like *smoke, drink and exercise* and relationships like *marriage and sister & brothers*.

The graph Fig. 1 showed the intensive overlapped area from 0.3 to 0.8 in the P-score curve. There is only one sample out of the overlap area. Therefore, the weak common support problem would not happen and confuse the regression.

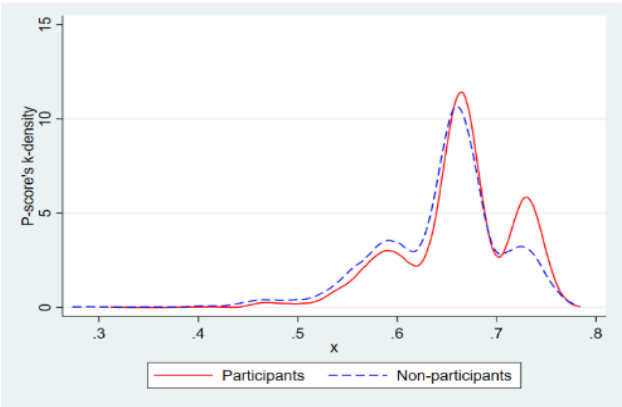


Figure 1. P-score distribution on two groups

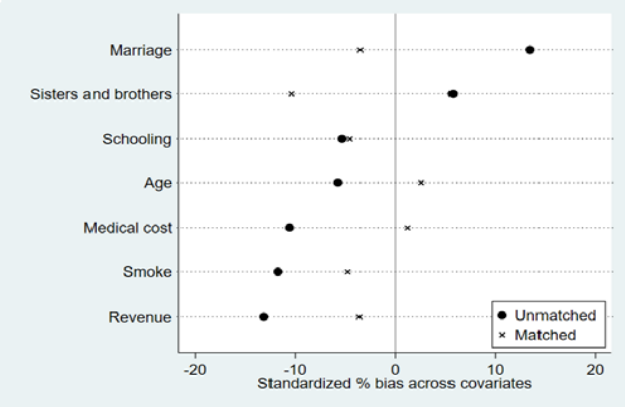


Figure 2. P-score bias reduction by kernel matching

There was much evidence about a fully efficient estimator could be yielded by a weighted least squares regression which weight the control observations according to their propensity score [22]. In this case, the study followed the general method of hackman's two-stage evaluation in a job program, employing weighted estimator based on the matched baseline in DID function. The graph Fig. 2 showed the status in bias reduction where *marriage, age, medical cost, smoke* and *revenue* had a great improvement on matching, which provided an ideal preprocessing for further analysis.

From Table 3, column (1) is the result in section2. The lowest overlapped samples in the density curve were dropped gradually from column (2) to column (5). After the revision of PSM for samples, the adverse effect of NCMS on *SRH* and *CSRH* are fewer and even reverse the coefficient in column

(5) but all insignificant. In contrast, the regression outcome in ADL fortified the evidence in DID work. The average treatment effect increased significantly from 3% to 4%. For Panel B, the effect on cure on time is significantly strengthened. Unmatched DID results showed a 15% increase in medical cost when one enrolled, while with the kernel matching in the baseline, the positive effect was significantly fewer but still keep 9% in the minimum from the table.

Table 3. PSM-DID regression

	DID	DID with PSM			
Kernel match	No	Yes	Yes	Yes	Yes
Sample Trim	0%	0%	10%	20%	30%
	(1)	(2)	(3)	(4)	(5)
Panel A: Health outcomes					
SRH Self-reported health	-0.04 (0.03) [0.16]	-0.02 (0.02) [-0.30]	-0.01 (0.02) [-0.76]	-0.02 (0.03) [-0.54]	-0.01 (0.03) [-0.68]
ADL Activities of daily living	0.03** (0.02) [0.03]	0.04*** (0.01) [0.00]	0.04*** (0.01) [0.00]	0.04*** (0.01) [0.00]	0.04*** (0.01) [0.00]
CSRH Comparative Self-reported health	-0.01 (0.02) [0.43]	-0.01 (0.01) [0.58]	0.00 (0.01) [0.77]	0.00 (0.02) [0.80]	0.01 (0.02) [0.49]
Panel B: Health access & service					
Log medical cost	0.15** (0.07) [0.03]	0.10* (0.05) [0.07]	0.09* (0.06) [0.10]	0.11* (0.06) [0.06]	0.15** (0.06) [0.02]
Cure on time	0.07*** (0.02) [0.00]	0.09*** (0.02) [0.00]	0.09*** (0.02) [0.00]	0.10*** (0.02) [0.00]	0.10*** (0.02) [0.00]
Sample size	2554	2553	2387	2220	2053

Notes: Column (1) kept the results from column (5) in Table2. Column (2) PSMDD includes all samples, weighting control groups by their propensity score. Column (3) (4) (5) are still the kernel match but trim the baseline sample whose p-score density is the lowest for common support.

Although *SRH* was showed no improvement because of the enrollment, the negative coefficients are weakened. It meant the adverse selection existed in the NCMS enrollment. The treatment itself could never ignore the outcome, which means if there are some of the people want to assort the NCMS to improve their health status, or, their decision to enroll the treatment is based on the previous health status, then the endogenous question would be furthermore complicated than this article.

3.3. Lasso Regression in a DID Framework

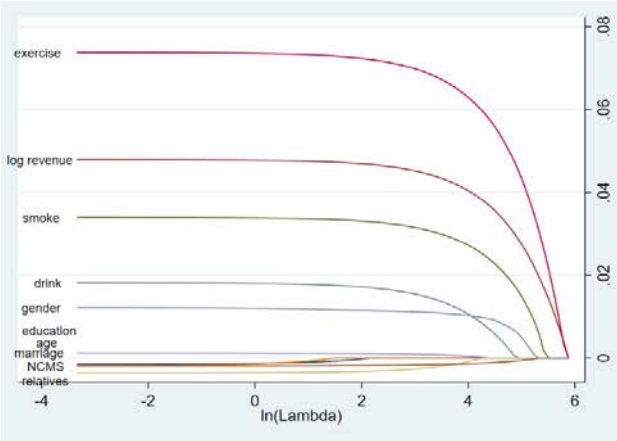
When the lots of data is available, non-linear methods may perform better than linear regression for the prediction problem [23]. Because of the existence of many instruments in the regression, some high-dimension problems could arise such as multicollinearity. Lasso (least absolute shrinkage and selection operator) regression could partly solve these by setting some variables to be exactly zero with the penalty term. In this case, the data from questionnaire was mostly dummy variables so that the specificity of each variable was unclear. Therefore, lasso regression was employed in this study for better variables in the OLS estimation through the selection. The minimization problem for lasso regression could be described as in Eq. 2:

$$\frac{1}{N}(y_i - x_i'\beta)^2 + \frac{\lambda}{N}\|\psi\beta\|_1 \tag{2}$$

Where λ denotes the penalty level. By introducing penalty term $\frac{\lambda}{N}\|\psi\beta\|_1$ in the residual sum of squares (RSS), overflowed variables were cleared automatically. The graph Fig. 3 illustrated the process of regularization by the variance of λ . Y-axis in the graph represent the estimated coefficient of each variable.

In this study, 10-fold cross-validation was used to find a best penalty parameter λ . λ value that minimizes the mean-squared prediction error (MSPE) was selected as the one in the lasso regression next.

The real line in the Fig. 4 indicated the value of λ and its MSPE. The dashed line indicated the largest λ at which the MSPE is within one standard error of the minimal MSPE but not used in this study.



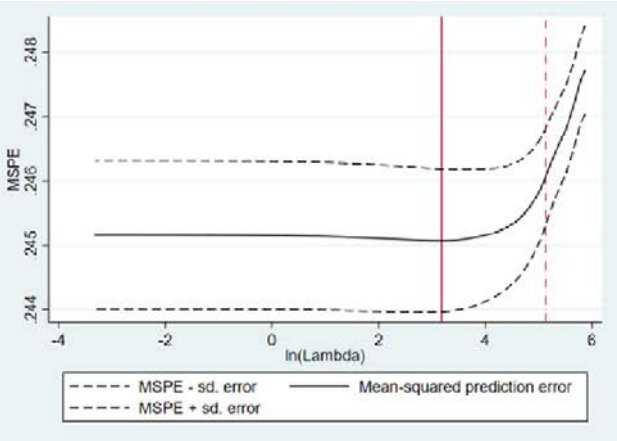


Figure 3. Coefficient path in lasso against SRH **Figure 4**. MSPE path in lasso against SRH

The expansion of regularized regression in this experiment was important because the estimation results in 3.1 and 3.2 may have the same variable problem. Especially, lasso has great advantage for interpretation when small number of the effects was moderate-sized within the regression [24]. In this case, Lasso could modify the causality in the moderate model like this study. Therefore, the similarity and difference between lasso result and previous results were valuable to observe and discuss. In this study, the variable selection was employed based on residual dependents where all fixed effects in the previous framework 3.1 worked.

Table 4. Lasso regression and Post-estimation

	lambda	OLS	Lasso	Post-OLS
Variable selection		No	Yes	Yes
	(1)	(2)	(3)	(4)
SRH	24.08	-0.04	-	-
ADL	7.04	0.03**	0.02***	0.02***
CSRH	13.59	-0.01	-	-
Log medical cost	155.02	0.15**	0.00	0.03
Cure on time	31.86	0.07***	0.02***	0.02***
Sample size	-	2554	2554	2554

Notes: Column (1) represented the best lambdas against five health outcomes from 10-fold cross-validation. Column(2) kept the results without modifying from column (5) in Table2, section3.1. Column (3) indicated the results of regularized regression. Column (4) is the results from a two-stage regression which resorts to lasso to select the suitable variables. horizontal line – indicated the NCMS effect was dropped in lasso. All the commands used in Stata could be found in the website here https://statalasso.github.io.

From Table 4, the difference in the regression results between column (2) and column (3)(4) was obvious except *ADL*. In terms of direct health outcomes, *SRH* and *CSRH*, the coefficients of NCMS were dropped which meant the NCMS enrollment indication may unclear within these two regressions. The *medical cost* effect varied most in all five indicators which decreased from 0.15 to just zero in lasso regression. Likewise, the *cure on time* effect fall down to 0.02 in both variable- modified regressions. The reason of this significant effect in medical & cure outcomes maybe the interactions in the variables before. Only *ADL* kept its moderate effect within the range of 0.02-0.03.

4. Conclusions

Using 2005-2008 datasets from the CLHLS, the study investigated the health-related benefit on rural elders who had enrolled the national health insurance NCMS. This study employed DID at first to difference out the time unvarying unobservable. Also, I used other estimations includes DID with kernel match and lasso regression, attempting to correct the omitted variable bias and address the self-selection problem within the experiment. Following the identifying assumption of common trend, no evidence showed a significant health effects of NCMS, which was consistent with some literatures [25]. But a significantly 15% upwards in their out-of-pockets medical cost.

Limited evidence for the association between health-related outcomes and enrolment were showed in DID estimation. DID is a flexible framework of causal inference concerning its advantage to combine with some other procedures. Therefore, a two-stage weighted regression was employed. However, with the matched and relatively more similar samples, the results were still not salient enough to indicate the positive effect of NCMS. Although the out of pocket cost decreased significantly by trimming the low-similar samples, there were no improvement in releasing the medical burden in participants. There are some literatures that said the self-reported maybe not enough representative than that of relative figure, but no significant coefficient showed for comparative change of health status. Therefore, the measurement bias still not be concluded.

To show further analysis in the study, this work associated with the lasso regression. The results indicated the variable selection problem in the previous model. The interplay effect between dummy variables was considerable. It was testified that the NCMS enrollment was unclear as an indicator for the measurement in *SRH* and *CSRH*. For direct health measurement, more detailed health data should be considered in the model calculation for clearer interpretation. In addition, the significant change in some variables indicated that the necessity to select variables carefully. Especially when the results were moderate, the effect could be varied much for modified variables.

Although the data shows relatively evident for the improvement in physical activity capability and the medical facility , it must be acknowledged that several defects limited the accuracy of the study. First, the results from PSM-DID varied from different kernel types and sample trimming. Because of the flexibility of the preprocessing method, the matched effect in the baseline still need further discussion. Second, those unobserved time-varying factors could arise confounding influence for the parallel assumption of DID even though the matching preprocess in DID model could help. Third, it may be uncertain for the precision of measurement in the questionnaire provided by answers from rural elders. The varied outcomes from the lasso regression indicated the necessity of accurate measurement in a proper scale. For the dependents, more specific indicators of objective or physician- assessed health status could do better than self-reported measures of health, like the steady effect of *ADL* in this study. Finally, the availability of micro-data on objective health outcomes would be an invaluable tool for more precise public health assessments of the insurance effects in the future.

5. References

[1] Khandker, S., B. Koolwal, G., & Samad, H. (2009). Handbook on impact evaluation: quantitative methods and practices: The World Bank.

[2] Hadley, J., & Waidmann, T. J. H. S. R. (2006). Health insurance and health at age 65: implications for medical care spending on new Medicare beneficiaries. 41(2), 429-451.

[3] Gakidou, E., Lozano, R., González-Pier, E., Abbott-Klafter, J., Barofsky, J. T., Bryson-Cahn, C., . . . Murray, C. J. J. T. L. (2006). Assessing the effect of the 2001–06 Mexican health reform: an interim report card. 368(9550), 1920-1935.

[4] Sepehri, A., Sarma, S., & Simpson, W. J. H. e. (2006). Does non - profit health insurance reduce financial burden? Evidence from the Vietnam living standards survey panel. 15(6), 603-616.

[5] Wagstaff, A., Lindelow, M., Jun, G., Ling, X., & Juncheng, Q. (2007). Extending health insurance to

the rural population: an impact evaluation of China's new cooperative medical scheme: The World Bank.

[6] Lei, X., & Lin, W. J. H. e. (2009). The new cooperative medical scheme in rural China: Does more coverage mean more service and better health? , 18(S2), S25-S46.

[7] Cheng, L., Liu, H., Zhang, Y., Shen, K., & Zeng, Y. J. H. e. (2015). The impact of health insurance on health outcomes and spending of the elderly: evidence from China's new cooperative medical scheme. 24(6), 672-691.

[8] Strauss, J., Lei, X., Park, A., Shen, Y., Smith, J. P., Yang, Z., & Zhao, Y. J. J. o. p. a. (2010). Health outcomes and socio-economic status among the elderly in China: Evidence from the CHARLS Pilot. 3(3-4), 111-142.

[9] Xian-xin, Z. J. P., & Economics. (2010). A Comprehensive Analysis of the Health of China's Elderly Population [J]. 5, 014.

[10] Data collected on opendata.pku.edu.cn/.

[11] https://sites.duke.edu/centerforaging/programs/chinese-longitudinal-healthy-longevity-survey- clhls/.

[12] Gu, D. (2007). General data quality assessment for the 2005 CLHLS wave. Retrieved from

[13] Yi, Z. (2008). Introduction to the chinese longitudinal healthy longevity survey (CLHLS). In Healthy longevity in China (pp. 23-38): Springer.

[14] Schütte, S., Chastang, J.-F., Parent-Thirion, A., Vermeylen, G., & Niedhammer, I. J. S. j. o. p. h. (2013). Social differences in self-reported health among men and women in 31 countries in Europe. 41(1), 51-57.

[15] DeSalvo, K. B., Bloser, N., Reynolds, K., He, J., & Muntner, P. J. J. o. g. i. m. (2006). Mortality prediction with a single general self-rated health question. 21(3), 267.

[16] Huisman, M., Deeg, D. J. J. S. s., & medicine. (2010). A commentary on Marja Jylhä's “What is self-rated health and why does it predict mortality? Towards a unified conceptual model”(69: 3, 2009, 307–316). 70(5), 652-654.

[17] Katz, S., Downs, T. D., Cash, H. R., & Grotz, R. C. J. T. g. (1970). Progress in development of the index of ADL. 10(1_Part_1), 20-30.

[18] Wang, H., Zhang, L., Yip, W., Hsiao, W. J. S. s., & medicine. (2006). Adverse selection in a voluntary Rural Mutual Health Care health insurance scheme in China. 63(5), 1236-1245.

[19] Angrist, J. D., & Pischke, J.-S. (2008). Mostly harmless econometrics: An empiricist's companion: Princeton university press

[20] Rosenbaum, P. R., & Rubin, D. B. J. B. (1983). The central role of the propensity score in observational studies for causal effects. 70(1), 41-55.

[21] Heckman, J. J., Ichimura, H., & Todd, P. E. J. T. r. o. e. s. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. 64(4), 605-654.

[22] Hirano, K., Imbens, G. W., & Ridder, G. J. E. (2003). Efficient estimation of average treatment effects using the estimated propensity score. 71(4), 1161-1189.

[23] Varian, H. R. J. J. o. E. P. (2014). Big data: New tricks for econometrics. 28(2), 3-28.

[24] Tibshirani, R. J. J. o. t. R. S. S. S. B. (1996). Regression shrinkage and selection via the lasso. 267-288.

[25] Chen, Y., & Jin, G. Z. J. J. o. h. e. (2012). Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. 31(1), 1-14.