

THE IMPACT OF HEALTH INSURANCE ON HEALTH OUTCOMES AND SPENDING OF THE ELDERLY: EVIDENCE FROM CHINA'S NEW COOPERATIVE MEDICAL SCHEME

LINGGUO CHENG^{a,*}, HONG LIU^{b,†}, YE ZHANG^a, KE SHEN^c and YI ZENG^{d,e}

^a*School of Business, Nanjing University, Nanjing, China*

^b*China Economics and Management Academy, Central University of Finance and Economics, Beijing, China*

^c*School of Social Development and Public Policy, Fudan University, Shanghai, China*

^d*Center for Healthy Aging and Development Studies, National School of Development, Peking University, Beijing, China*

^e*Center for the Study of Aging and Human Development, Geriatrics Division of School of Medicine, Duke University, Durham, NC, USA*

ABSTRACT

This paper investigates the effects of China's New Cooperative Medical Scheme (NCMS) on health outcomes and healthcare expenditure of the elderly in rural China, using panel data from the 2005 and 2008 waves of the Chinese Longitudinal Healthy Longevity Survey. We employ a strategy that combines propensity score matching with a difference-in-differences approach to address selection bias. Results show that the NCMS has significantly improved the elderly enrollees' activities of daily living and cognitive function but has not led to better self-assessed general health status. We find no significant effect of NCMS on mortality for the previously uninsured elderly in NCMS counties, although there is moderate evidence that it is associated with reduced mortality for the elderly enrollees. We also find that the elderly participants are more likely to get adequate medical services when sick, which provides a good explanation for the beneficial health effects of NCMS. However, there is no evidence that the NCMS has reduced their out-of-pocket spending. Furthermore, we also find that low-income seniors benefit more from NCMS participation in terms of health outcomes and perceived access to health care, suggesting that the NCMS helps reduce health inequalities among the rural elderly. Copyright © 2014 John Wiley & Sons, Ltd.

Received 28 December 2012; Revised 8 December 2013; Accepted 13 March 2014

KEY WORDS: China; health insurance; health outcome; health spending; the elderly

1. INTRODUCTION

Many developing countries are currently trying to expand health insurance coverage to improve health outcomes and to reduce illness-induced impoverishment. However, credible evidence that expanded insurance coverage improves health status and lowers out-of-pocket health expenditures is still lacking in the literature (Ekman, 2004). This paper aims at contributing to the literature by examining the impact of China's recent expansion of the rural public health insurance system, the New Cooperative Medical Scheme (NCMS), on health outcomes and spending. In particular, we focus on the elderly population in rural China, who are most vulnerable to ill-health and diseases.

From the 1950s through the 1970s, a commune-based healthcare financing system, known as the Cooperative Medical Scheme (CMS), provided healthcare coverage for most of the Chinese rural population (Feng *et al.*, 1995; You and Kobayashi, 2009). However, market-oriented economic reform launched in

*Correspondence to: School of Business, Nanjing University, 22 Hankou Road, Nanjing 210093, China. E-mail: chenglingguo@gmail.com

†Hong Liu (irisliu2000@gmail.com) contributed equally to this work with Lingguo Cheng.

1978 led to the dismantling of the CMS due to the cessation of financial support from the rural collective economy. Consequently, the health insurance coverage rate dropped from 90% before the reform to less than 10% in 1993 (Liu *et al.*, 1995; Wagstaff and Lindelow, 2008). Despite a number of local attempts to reestablish some form of rural CMS during the 1990s, nearly 80% of rural residents (about 640 million) did not have any health insurance in 2003 (Ministry of Health, 2004). Meanwhile, the cost of medical care has increased explosively in the market-oriented healthcare reform since the 1990s (Eggleston *et al.*, 2008; Yip and Hsiao, 2009). With limited insurance coverage, many rural families experienced decreasing availability, rising cost, and compromised quality of health care and were driven into or back into poverty as a result (Liu *et al.*, 1995; Tang *et al.*, 2008; Yip and Hsiao, 2009).¹ The situation was more severe for the rural elderly, who tended to be poorer, less educated, and at increased risk of health problems. It is estimated that about 44.4% of rural elderly had inadequate access to health care in 2003 (Ministry of Health, 2004).

To address these problems, China launched a nationwide project known as the NCMS in rural China in 2003. Different from the universal health insurance in most developed economies, the NCMS is a government-run voluntary insurance program operated at the county level, primarily focusing on the coverage of catastrophic diseases (State Council, 2002). Following the broad guidelines issued by the central government, local governments have considerable discretion over the design and implementation of their specific programs. As a result, the NCMS varies considerably across counties with respect to premiums and benefits (Brown *et al.*, 2009). Basically, all rural households in NCMS counties are eligible and can choose to participate in the program for either all or none of their household members. It is financed by individual contributions and heavy subsidies from central and local governments. During the period covered by the study (2005–2008), the minimum requirement for the household contribution was RMB 20 per person, and the subsidy payment was RMB 40 per person. In 2008, the standard subsidy level from central and local governments rose to RMB 80 per person. Because NCMS is intended for reimbursement mainly of catastrophic expenses, all counties cover inpatient care but vary considerably in the coverage of outpatient care and the reimbursement levels. Most counties cover outpatient services through a household account (about 65%) or on a pooling basis (about 7%), whereas the rest cover outpatient services only for catastrophic diseases (about 11%) or do not cover outpatient services at all (about 17%) (Du and Zhang, 2007; Lei and Lin, 2009; Wagstaff *et al.*, 2009a; You and Kobayashi, 2009; Chen and Jin, 2012).²

The NCMS began as a pilot program in 2003 and expanded rapidly: from 310 out of the total of 2861 rural counties in 2004, to 1451 counties in 2006, and to almost all counties in 2008. It provided health insurance coverage for about 96% (836 million) of the rural population in 2010 (NBS, 2011). The literature so far has found that the NCMS has improved access to outpatient and inpatient care but had a limited effect in reducing out-of-pocket spending burdens (Lei and Lin, 2009; Sun *et al.*, 2009, 2010; Wagstaff *et al.*, 2009a; Yip and Hsiao, 2009; Babiartz *et al.*, 2010, 2012). However, nearly 10 years after its inception, there is limited empirical research on how well the NCMS has improved the health of the rural population, which is central to the goals of China's healthcare reform.

This paper contributes to the literature by systematically evaluating the effects of the NCMS on health outcomes and spending of the rural elderly, using panel data from the Chinese Longitudinal Healthy Longevity Survey (CLHLS). It is a longitudinal survey project with five waves since 1998, mainly focusing on the elderly population, and includes multiple measures of health status as well as measures of perceived healthcare access and expenditure. We are interested in the impact of the NCMS on measurable outcomes of the insured, that is, the average treatment effect on the treated (ATT). To address potential selection bias due to the voluntary nature of the NCMS, we employ a propensity score matching with difference-in-differences (PSMDD) strategy, which has been commonly used in the literature on the impact of the NCMS (Lei and Lin, 2009; Wagstaff *et al.*, 2009a; Wang

¹It was reported that 33% of poor households identified illness or injury as the cause of their poverty (Ministry of Health, 2004).

²By 2007, an increasing number of counties had included outpatient services at different levels of healthcare providers (e.g., hospitals, township health centers, and village clinics) and physical checkups in the benefit package of the NCMS (Du and Zhang, 2007; Lei and Lin, 2009; Babiartz *et al.*, 2010).

et al., 2009; Chen and Jin, 2012). Furthermore, we carry out a falsification test to provide supportive evidence for the validity of our main strategy and also check the robustness of our results with different sample subgroups.

This study shows that the NCMS has significantly improved the elderly enrollees' activities of daily living (ADL) and cognitive function but has not led to better self-assessed general health status. We find no significant effect of the NCMS on 3-year mortality for the previously uninsured elderly in NCMS-exposed counties, although there is moderate evidence that NCMS is associated with reduced mortality for the elderly enrollees. We find no evidence that the NCMS has reduced the elderly enrollees' out-of-pocket spending. We also find that the elderly participants are more likely to get adequate medical services when sick, which provides a good explanation for the beneficial health effects of the NCMS. Furthermore, it is found that low-income seniors benefit more from NCMS participation in terms of health outcomes and perceived access to health care.

The rest of the paper is organized as follows. Section 2 reviews some of the related empirical literature. Sections 3 and 4 describe the data and empirical strategy. Section 5 presents the results of the empirical analyses. Section 6 concludes the paper with a brief discussion.

2. LITERATURE REVIEW

There is extensive literature on the influence of health insurance systems in developed countries. With some exceptions (Hadley, 2003), considerable evidence has shown that health insurance coverage is associated with better self-reported health status and lower mortality risk (Franks *et al.*, 1993; McWilliams *et al.*, 2004; Hadley and Waidmann, 2006). These findings are limited, however, by concerns about selection biases inherent in these observational studies. Recently, some literature has explored the causal relationship between health insurance and a number of outcomes by using expansion of public health insurance as a natural experiment and yielded mixed results. For example, studies show that the expansion of Medicaid coverage in the USA has been effective in improving mothers' access to prenatal care and children's utilization of preventive care, reducing the incidence of low birth weight and infant mortality, but had no beneficial effect on the health status of older children (Currie and Gruber, 1996a, 1996b, 1997; Currie *et al.*, 2008). Literature focusing on the US elderly population also shows that the increase of health insurance coverage for the elderly at age 65 by Medicare has led to more utilization of medical care, better self-reported health status, and lower out-of-pocket expenses but has little effect on mortality (Decker and Remler, 2004; Card *et al.*, 2008; Finkelstein and McKnight, 2008). Studies on Taiwan find that the introduction of the National Health Insurance program has reduced mortality for infants (Chou *et al.*, 2011), but the results for the elderly are mixed (Chen *et al.*, 2007; Chang, 2012; Keng and Sheu, 2013).

Similar studies are limited for developing countries, many of which have been attempting to establish or expand their own public health insurance programs recently, as China is doing. Existing literature investigates the impact of public health insurance in Vietnam (Jowett *et al.*, 2004; Sepehri *et al.*, 2006), Colombia (Panopoulou and Velez, 2001; Trujillo *et al.*, 2005), and Mexico (Gakidou *et al.*, 2006). Moreover, there are two previous studies on the effects of local attempts to resurrect community health insurance in rural China before the introduction of the NCMS. One is that of Wang *et al.* (2009), which shows that the community-based health insurance scheme improved the health status of participants in one of China's western provinces. The other study, by Wagstaff and Yu (2007), evaluates a reform project on both the demand and supply sides of the rural health sector in Gansu province of China and finds that it relieved the financial burden on individual households but had no effect on health care utilization and mixed effects on health outcomes.

A number of studies have been carried out to evaluate the impact of the NCMS since its inception in 2003. Most of the existing literature has shown that the NCMS significantly improved the enrollees' utilization of outpatient services and inpatient services (Wagstaff *et al.*, 2009a, 2009b; Babiarz *et al.*, 2010; Yu *et al.*, 2010; Babiarz *et al.*, 2012). Lei and Lin (2009) find that the NCMS increased the use of preventive care for the participants but had no significant effect on the utilization of formal medical care. Moreover, previous studies have provided mixed evidence of NCMS's impacts on health expenditures, probably because they study different stages of the program. Lei and Lin (2009) and Wagstaff *et al.* (2009a) find no evidence that the NCMS reduced

out-of-pocket expenditure during the pilot period of the NCMS, whereas Babiarz *et al.* (2010, 2012) find that enrollment in NCMS is associated with a decrease in out-of-pocket medical spending with the progress of the NCMS program. However, probably because of constrained funding and a low reimbursement rate (Yi *et al.*, 2009), financial protection provided by the NCMS is still limited, as most studies only find a modest effect of NCMS on the incidence of catastrophic health spending and on poverty due to health spending (Sun *et al.*, 2009, 2010; Yip and Hsiao, 2009; Babiarz *et al.*, 2010, 2012).

So far, there are only two formal studies that have considered the impact of the NCMS on health outcomes. Using panel data from the China Health and Nutrition Survey (2000–2006), Lei and Lin (2009) find that the NCMS had no impact on improving the enrollees' health status measured by self-reported health and sickness or injuries in the preceding 4 weeks; however, they also express concern about the data limitation in health measurement. Chen and Jin (2012) use data from the 2006 China Agricultural Census and also find that the NCMS has not affected the mortality rate of pregnant women and young children at the village level. They conclude that the positive association between the NCMS and a lower mortality rate is mainly driven by the endogenous introduction of the NCMS. Both studies employ PSMDD methods to address the selection bias, as we do in this paper; however, their empirical analyses mainly cover the piloting period of the NCMS and suffer from limited measures of health outcomes.

It is reported that the design and implementation of the NCMS were significantly improved from 2004 to 2007, as a result of increased government subsidies (Yi *et al.*, 2009). Taking into account this new evidence, it is necessary to reevaluate the earlier conclusions about the effectiveness of the NCMS, especially its health effects. In this paper, we contribute to the literature by examining the impact of the NCMS on health outcomes and healthcare costs during the period 2005–2008 and by identifying the beneficial health effects of the NCMS among rural Chinese elderly.

3. DATA AND SUMMARY STATISTICS

3.1. Data

Our data were drawn from the CLHLS, collected by the Center for Healthy Aging and Development Studies at Peking University and co-sponsored by the US National Institute on Aging. This survey was conducted in a randomly selected half of the counties and cities in 22 provinces (out of total 31 provinces),³ covering 85% of the total population in China. The CLHLS has six waves to date (1998, 2000, 2002, 2005, 2008, and 2011). It surveyed only the oldest old (aged 80 years and older) before 2002, but since 2002, younger elders (aged 60–79) have also been included in the project. The survey combines an in-home interview and a basic physical examination. A detailed description of the sampling design and data quality assessment of the CLHLS can be found elsewhere (Zeng *et al.*, 2001, 2002; Gu and Zeng, 2004; Gu, 2008). Extensive information was collected on demographic characteristics, family and household characteristics, lifestyles, diet, psychological characteristics, health, disability, socioeconomic conditions, and so on. However, the CLHLS has asked respondents about their health insurance coverage and healthcare expenditure only since 2005. For the purpose of this study, we mainly use the 2005 and 2008 waves of CLHLS data.⁴

3.2. Health indicators

We use multiple indicators for the health status of the elderly, including mortality within 3 years, self-reported health, and measured health. The CLHLS collected mortality data for the respondents who died between waves (i.e., 2005–2008 and 2008–2011) in interviews with a close family member of the deceased. The binary outcome variable *three-year mortality* indicates whether the survey respondent died during the later 3 years.

³The sample provinces include 10 eastern provinces, 8 central provinces, and 4 western provinces.

⁴The most recent wave of the CLHLS data was collected in 2011–2012 and has not been released yet. So we are not able to include the 2011 wave of CLHLS data. However, we have access to the information whether the respondents in wave 2008 had deceased prior to 2011 or had been lost to follow-up in wave 2011, which helps us construct the mortality indicator for the respondents in wave 2008.

Subjective health is measured by perceived general health, based on a four-item Likert scale survey question: 'How do you rate your health at present?' We construct a binary indicator of self-assessed health status, recoded as 'good' (excellent or good) versus 'poor' (fair or poor).

In addition to mortality, other objective health indicators include ADL, cognitive functions, incidence of severe illness, number of bedridden days, and measured hypertension. The survey questions about ADL and cognitive functions are based on international standards and adapted to the Chinese culture and social context with carefully conducted pilot study tests (Zeng *et al.*, 2002; Gu and Zeng, 2004). Specifically, ADL measures the physical function of the elderly, according to Katz's ADL index (Katz *et al.*, 1963). We construct a binary variable that equals 1 if the respondent reported no restriction in six daily activities—eating, dressing, moving, using the toilet, bathing, and continence—and equals 0 if any restriction in these activities was reported.

Following the literature (Folstein *et al.*, 1975; Crum *et al.*, 1993), cognitive functions are measured using the Mini-Mental State Examination (MMSE). The MMSE questionnaire includes 24 items regarding orientation, registration, attention, calculation, recall, and language, with a total score ranging from 0 to 30. In addition to the continuous specification of MMSE scores, we also use a binary variable indicating whether the respondent has good cognitive function, which is 1 if the MMSE score is no less than 24, and 0 otherwise (Folstein *et al.*, 1975).

The respondents were also asked about their incidence of severe diseases and resulting bedridden status in the preceding 2 years. We created a binary variable indicating whether the respondent has had serious illness in the preceding 2 years and a continuous variable equal to the number of bedridden days in the preceding 2 years.

In addition to these self-reported measures, the CLHLS data also contain information on blood pressure measured by the medical personnel during the face-to-face interviews. Following the medical standard, we define measured hypertension as having systolic blood pressure no less than 140 mmHg or diastolic pressure no less than 90 mmHg.

3.3. Health spending and perceived access to health care

There are two survey questions about individuals' healthcare expenditure in CLHLS. Respondents are asked, 'How much did you spend on medical costs last year?' and 'Of this, how much was paid by the family (self, spouse, children, etc.)?' Accordingly, we construct two outcome variables: total healthcare expenditure and out-of-pocket expenditure in the previous year, both measured in terms of 2005 prices.

The CLHLS did not gather direct information on healthcare utilization but asked respondents, 'Can you get adequate medical service when you are sick?' Those who answered 'no' were further asked, 'What's the primary reason that you didn't go to the hospital when it was necessary?' Based on the responses, we construct two binary variables measuring perceived access to health care, one indicating whether the respondent could get adequate medical services when sick and the other indicating whether the respondent had failed to get necessary care owing to costs.

3.4. Sample definition and descriptive statistics

Focusing on the subsample of rural elderly, we start with 8658 respondents who had been interviewed in wave 2005. We exclude the counties exposed to the NCMS before 2005⁵ and the respondents with any other types of health insurance during 2005–2008.⁶ These exclusions bring the sample down to 6683 rural respondents in 463 counties. Of these, 3299 (49.3%) survived and were re-interviewed in 2008, 2506 (37.5%) died before the 2008 survey, and the remaining 878 (13.1%) were lost to follow-up (Table AI in the online supporting information).

⁵About 14% of the total respondents were from counties that had implemented the NCMS before 2005. As the 2002 CLHLS survey did not collect the information on individual health insurance status and healthcare expenditure, we cannot include this sample in our pre-post treatment-control study. Assuming that the respondents were all uninsured in wave 2002, we can use CLHLS 2002–2005 to check the robustness of our main estimates. Although unreported here, the results show that the estimates of NCMS impact are robust to the inclusion of those experimental counties between 2002 and 2005.

⁶Approximately 12% of the respondents had other types of health insurance during 2005–2008, such as commercial health insurance, government employee insurance, urban employee medical insurance, or other health insurance.

Panel attrition may be a source of bias in the present study if the loss to follow-up is not random and the variables affecting attrition are correlated with our outcome variables. To address this concern, we investigate whether those lost to follow-up are different from the follow-up sample in two ways, based on characteristics in the base year of 2005. First, we compare pre-attrition characteristics by attrition status in Appendix Table AII in the online supporting information. We find no significant difference in health outcomes, healthcare access, or spending between the attriting and nonattriting samples. However, those who were lost to follow-up in 2008 tended to have higher education levels and incomes and were more likely to have white-collar jobs before age 60. Second, to test statistically for attrition due to observables, we follow the method suggested in the literature (Beckett *et al.*, 1988; Ding and Lehrer, 2010; Keng and Sheu, 2013) and regress the baseline outcomes on baseline individual characteristics and their interactions with a binary indicator for attrition (i.e., loss to follow-up), using the full sample of wave 2005. The results presented in Appendix Table AIII in the online supporting information show that the coefficients on the attrition indicator and all of its interaction terms are jointly insignificant, indicating that the elderly lost to follow-up in wave 2008 are not systematically different from the remaining sample in health outcomes and spending.⁷ Thus, we can safely exclude the respondents lost to follow-up in wave 2008, and the potential attrition bias need not be a concern in this study.

In our study sample, no county was exposed to the NCMS in wave 2005, and almost all counties had implemented the program by 2008. To estimate the average treatment effect of enrolling in NCMS, we define the treatment and control groups mainly based on individual enrollment status in wave 2008. In a pre–post treatment–control study design for most outcomes except that for mortality, we further exclude the respondents deceased prior to 2008 (we return to this attrition issue in Section 5) and construct a balanced panel data set. The final sample consists of 3299 rural respondents in each wave. Of these, 2252 respondents, who had no health insurance in 2005 but participated in the NCMS in 2008, are classified as the treated group, and 1047 respondents, who had been uninsured for the entire period 2005–2008, are classified as the control group.

We also control for demographic and socioeconomic covariates that may affect enrollment in the NCMS, as well as health outcomes and spending, including age, gender, marital status (1 = married), years of schooling, occupational category before age 60 (1 = had a white-collar job, i.e., professional, managerial, or administrative), household income per capita, number of adult children alive at interview, co-residence with adult children, and health behavior (1 = exercise regularly). Geographic regional variables are also included in most specifications.

Table I provides a description of our panel sample and presents the outcomes and characteristics of the treated and control groups. Compared with the control group, the elderly in the treatment group had slightly better cognitive functions and were less likely to have measured hypertension in the pretreatment year. Although the respondents in both groups experienced some health deterioration due to aging between 2005 and 2008, the treatment sample had better health status than the control sample in terms of ADL, cognitive functions, bedridden days, and measured hypertension in the posttreatment year. Consistently, the treated group was more likely to get adequate medical services when sick, was less likely to report financial barriers to healthcare access, and had less out-of-pocket medical expenditure in 2008. Looking at the baseline characteristics, we can see that the elderly enrollees were younger, were more likely to be married, and had lower income and more adult children alive at interview than the nonenrollees.

4. EMPIRICAL FRAMEWORK

4.1. Method

We seek to identify the effect of the NCMS on health outcomes and spending of the elderly and focus on the ATT. Because take-up of the NCMS is voluntary, we need to deal with the selection bias due to unobserved

⁷Gu and Zeng (2004) and Gu *et al.* (2007) suggest that dropout attrition (i.e., loss to follow-up) in the CLHLS is generally lower than that in some panel surveys of older persons in the USA (e.g., Mihelic and Crimmins, 1997), and the potential panel attrition bias can be safely ignored.

Table I. Descriptive statistics by wave and treatment status

		Pretreatment year (wave 2005)		Posttreatment year (wave 2008)	
Variables	Full sample (<i>N</i> = 6598)	Treated (<i>N</i> = 2252)	Control (<i>N</i> = 1047)	Treated (<i>N</i> = 2252)	Control (<i>N</i> = 1047)
Panel 1: health outcomes					
Self-reported health is good	0.48 (0.50)	0.52 (0.50)	0.51 (0.50)	0.44 (0.50)	0.45 (0.50)
No ADL limitation	0.87 (0.34)	0.89 (0.31)	0.89 (0.31)	0.86*** (0.35)	0.82 (0.38)
MMSE score	22.89 (8.60)	24.45*** (7.13)	23.47 (7.85)	22.17*** (9.27)	20.51 (9.87)
Cognitive function is good	0.65 (0.48)	0.72*** (0.45)	0.68 (0.47)	0.63*** (0.48)	0.54 (0.50)
Had serious illness in the preceding 2 years	0.15 (0.36)	0.15 (0.35)	0.16 (0.37)	0.16 (0.37)	0.15 (0.36)
Bedridden days in the preceding 2 years	6.84 (42.64)	7.21 (46.30)	6.37 (31.32)	5.74* (36.13)	8.91 (55.29)
Measured hypertension at interview	0.39 (0.49)	0.40* (0.49)	0.43 (0.50)	0.35*** (0.48)	0.42 (0.49)
Panel 2: perceived access to health care					
Get adequate medical care when sick	0.87 (0.33)	0.85 (0.35)	0.85 (0.36)	0.91*** (0.29)	0.85 (0.36)
Fail to get necessary care due to costs	0.08 (0.27)	0.10 (0.29)	0.10 (0.29)	0.05*** (0.22)	0.07 (0.26)
Panel 3: health spending (deflated to the 2005 price level, RMB)					
Total medical expenditure	872.36 (2409.61)	694.34 (1972.18)	763.36 (1839.01)	985.34 (2675.74)	1125.71 (3056.90)
Out-of-pocket medical expenditure	815.47 (2242.14)	673.67 (1929.05)	746.87 (1831.17)	862.81** (2293.85)	1094.50 (2996.69)
Panel 4: demographic and socioeconomic characteristics					
Age	83.05 (11.05)	81.20** (10.99)	82.03 (10.75)	84.34** (11.00)	85.22 (10.80)
Male	0.42 (0.49)	0.42 (0.49)	0.43 (0.50)	0.42 (0.49)	0.43 (0.49)
Married	0.35 (0.48)	0.40*** (0.49)	0.33 (0.47)	0.33** (0.47)	0.30 (0.46)
Years of schooling	1.39 (2.51)	1.39 (2.46)	1.39 (2.63)	1.39 (2.46)	1.39 (2.63)
Had a white-collar job before age 60	0.02 (0.15)	0.02* (0.14)	0.03 (0.16)	0.02* (0.14)	0.03 (0.16)
Household income per capita	8994.61 (21,029.08)	9087.27*** (23,226.14)	14,640.80 (29,555.50)	6291.42*** (13,457.46)	8982.67 (17,861.07)
Exercising regularly	0.24 (0.43)	0.25 (0.43)	0.22 (0.42)	0.24 (0.42)	0.25 (0.44)
No. of adult children	3.89 (1.88)	4.01*** (1.90)	3.63 (1.91)	3.98*** (1.85)	3.68 (1.81)
Co-reside with adult children	0.58 (0.49)	0.59 (0.49)	0.57 (0.50)	0.60*** (0.49)	0.54 (0.50)

t-test was applied for pairwise comparisons of the treated and control groups in each wave.

****p* < 0.01;

***p* < 0.05;

**p* < 0.1.

heterogeneity between enrollees and nonenrollees. Our main empirical strategy is to use PSMD (for most outcomes except mortality within 3 years) to remove selection on observables and unobservables that are constant over time or have a common time trend between the treated and control groups (Heckman *et al.*, 1998, 1999; Imbens, 2004; Wagstaff *et al.*, 2009a; Wang *et al.*, 2009). PSMD compares differences between the pretreatment and posttreatment outcomes of participants with those of nonparticipants.

Ideally, the average treatment effect on the treated would be estimated as

$$ATT = E\left(Y_{i,\text{post}}^P - Y_{i,\text{pre}}^P | X_i^P, Z_i^P, D_i = 1\right) - E\left(Y_{i,\text{post}}^{NP} - Y_{i,\text{pre}}^{NP} | X_i^P, Z_i^P, D_i = 1\right) \quad (1)$$

where Y_i^P and Y_i^{NP} represent, respectively, the treatment and nontreatment outcomes of participant i before and after the treatment, as denoted by the second subscripts (post and pre); X_i^P is a set of observed individual characteristics of participant i ; Z_i^P indicates individual-specific unobservable attributes of participant i ; and D_i is a binary variable indicating whether individual i participates in the NCMS. Because $E\left(Y_{i,\text{post}}^{NP} - Y_{i,\text{pre}}^{NP} | X_i^P, Z_i^P, D_i = 1\right)$ is unobserved, it is typically assumed in the matching literature that participation status can be treated as random if the treated and control groups are matched on observables $X^P = X^{NP} = X$ (Rosenbaum and Rubin, 1985), so that $E\left(Y_i^{NP} | X_i, Z_i^P, D_i = 1\right) = E\left(Y_i^{NP} | X_i, Z_i^{NP}, D_i = 0\right)$. We have

$$ATT = E\left(Y_{i,\text{post}}^P - Y_{i,\text{pre}}^P | X_i, Z_i^P, D_i = 1\right) - E\left(Y_{i,\text{post}}^{NP} - Y_{i,\text{pre}}^{NP} | X_i, Z_i^{NP}, D_i = 0\right) \quad (2)$$

Moreover, if Z_i is time invariant or is time variant but has the same time trend between participants and non-participants, the difference in the Z_i can be eliminated by taking double differences. On the basis of the conditional independence assumption that the potential outcomes are independent of participation status conditional on the covariates X (Rosenbaum and Rubin, 1983), we can estimate the ATT conditional on the propensity score, $P(D_i = 1 | X_i)$. Then Equation 2 can be rewritten as

$$ATT = E\left(Y_{i,\text{post}}^P - Y_{i,\text{pre}}^P | P(D_i = 1 | X_i), D_i = 1\right) - E\left(Y_{i,\text{post}}^{NP} - Y_{i,\text{pre}}^{NP} | P(D_i = 1 | X_i), D_i = 0\right) \quad (3)$$

To implement PSMDD, we first estimate a probit model to obtain the propensity score—that is, the probability of being in the treatment group given the set of baseline observable characteristics X_i . The outcome variable is a binary variable indicating whether the individual is in the treatment group or the control group. The covariates include age, age squared, gender, marital status, education, occupation before age 60, household income, familial support (measured by co-residence with adult children), health behavior, self-rated health, perceived healthcare access, and regional dummies.⁸ With the estimated propensity score, we restrict the analyses to individuals in the common support. We match each treated individual with one or more individuals in the control group that are close to him or her in propensity score. Specifically, we adopt kernel matching algebra⁹ and use the weighted average of all comparable individuals in the control group to construct the counterfactual outcome for each treated individual. The weight assignment depends on the choice of kernel function and bandwidth parameters. In this paper, we use the Gaussian kernel with bandwidth of 0.06. We also try other bandwidths (0.1, 0.08, 0.04, and 0.02) to check the sensitivity of our results to the choice of bandwidth.

4.2. Matching and balancing tests

Figure 1 shows the histograms of the estimated propensity scores for the treated and control groups before matching. It is evident that the distributions of propensity score for the two groups overlap sufficiently so that we can perform matching over the region of common support.¹⁰

Table II reports the balancing test results based on Gaussian kernel matching. The upper part shows the balancing property for each observed covariate between the treated and control groups and the reduction in bias achieved through matching. We first report the means of each characteristic for the treated and control groups

⁸Following Wagstaff *et al.* (2009a), we do not force matching within the same province, which might lead to poor-quality matching because the sample distribution is highly dispersed across provinces.

⁹Compared with other matching algorithms, kernel matching has one major advantage of having a lower variance because more information is used. In addition, Frölich (2004) also suggests that, over a range of data-generating processes, kernel matching consistently does well on a mean-squared-error criterion.

¹⁰Only eight observations lying off the common support are excluded.

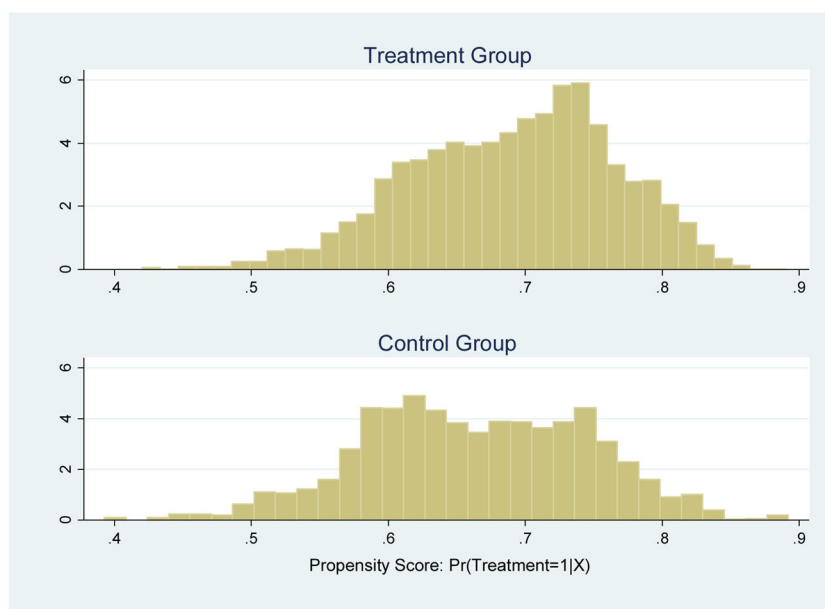


Figure 1. Propensity score histograms

and then check the significance of the difference in means by the two-sample t -test before and after matching, respectively. Columns (5) and (10) report the standardized difference in means (so-called *bias*) before and after matching, and the last column reports the percentage reduction in the bias through matching.¹¹ The results suggest that by the matching, almost all observable characteristics are balanced well between the treated and control groups. The initial differences in the two groups are reduced considerably and become statistically insignificant (at the 5% level).

In the second part of Table II, we check the overall balancing properties of the matching, by comparing the joint significance of all the matching variables in the probit models before and after matching. The first column shows that the pseudo- R^2 is much lower for the matched sample than for the raw sample. As shown in the middle two columns, the matching covariates are jointly significant before matching, but insignificant with a p -value of 0.393 after matching, indicating that matching improves the overall balance. Consistently, the last two columns show that both mean and median of the absolute standardized bias have been reduced substantially by the matching.

5. EMPIRICAL RESULTS

5.1. Main effects of NCMS: PSMDD estimates

Table III reports the PSMDD estimates of NCMS impacts on self-reported and measured health outcomes, perceived healthcare access, and medical expenditures, using the balanced panel of 3299 elderly respondents. The main bandwidth parameter is 0.06, and the results are robust to the selection of alternative bandwidth parameters. Thus, we primarily discuss the main estimates with the common bandwidth 0.06.

¹¹The standardized bias is defined as the difference in the mean values of the treated and control groups, divided by the square root of the average sample variance in the treated and control groups (Rosenbaum and Rubin, 1985).

Table II. Balancing tests from kernel matching

Part A. Test the balancing property for each observed covariate											
Variables	Prematching					Postmatching					% reduction bias
	Control		<i>t</i>	<i>p</i> > <i>t</i>	% bias	Treated		Control	<i>t</i>	<i>p</i> > <i>t</i>	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	
											(11)
Age	80.78	81.61	-2.02	0.043	-7.8	80.78	81.42	0.15	0.881	-6	22.5
Age squared	66.41	67.74	-1.93	0.053	-7.4	66.41	67.45	0.15	0.884	-5.8	21.8
Male	0.43	0.43	-0.14	0.892	-0.5	0.43	0.42	0.23	0.817	1.4	-162.1
Had a white-collar job before age 60	0.02	0.03	-1.83	0.067	-6.8	0.02	0.02	1.27	0.203	2	71.1
Years of schooling	1.42	1.41	0.11	0.915	0.4	1.42	1.37	0.37	0.714	2.1	-415.2
Married	0.41	0.33	4.07	0	15.7	0.41	0.37	-1.14	0.255	7.6	51.9
Co-residing with adult children	0.58	0.57	0.9	0.369	3.4	0.58	0.58	-0.63	0.53	-0.3	90.3
Self-reported health is good	0.52	0.51	0.53	0.599	2	0.52	0.53	-0.48	0.633	-0.7	66.1
Getting adequate medical care when sick	0.86	0.85	1.3	0.194	4.9	0.86	0.86	-0.41	0.684	0.9	82
Log(1 + household income per capita)	7.60	7.93	-4.76	0	-17.7	7.60	7.68	1.24	0.214	-4.5	74.5
Exercising regularly	0.25	0.22	1.57	0.117	6.1	0.25	0.25	-0.92	0.36	0	99.9
Central provinces	0.25	0.33	-4.57	0	-17.3	0.25	0.27	1.56	0.118	-5	71.3
Western provinces	0.46	0.34	6.42	0	24.9	0.46	0.41	-1.87	0.061	9.5	61.8
Part B. Test the overall balance											
Sample	Pseudo <i>R</i> ²	LR χ^2				<i>p</i> > χ^2				Mean bias	Median bias
Raw sample before matching	0.024	93.54				0.000				8.8	6.8
Matched sample after matching	0.004	13.73				0.393				3.5	2.1

All results are computed using the Stata module of `psmatch2` (Leuven and Sianesi, 2012).

All results are computed using the Stata module of psmatch2 (Leuven and Sianesi, 2012).

Table III. PSMDD estimates of NCMS impact

	Bandwidth				
	Main estimates	Sensitivity of results to choice of bandwidth			
	0.06	0.1	0.08	0.04	0.02
Panel 1: health outcomes					
Self-reported health is good	−0.005 (0.030)	−0.009 (0.029)	−0.008 (0.030)	−0.003 (0.030)	−0.003 (0.031)
Observations	2904	2904	2904	2904	2904
No ADL limitation	0.038** (0.019)	0.037** (0.018)	0.038** (0.018)	0.038** (0.019)	0.037** (0.019)
Observations	3159	3159	3159	3159	3159
MMSE score	0.932** (0.440)	0.947** (0.426)	0.943** (0.432)	0.903** (0.447)	0.853* (0.447)
Observations	3159	3159	3159	3159	3159
Cognitive function is good	0.035 (0.022)	0.037* (0.022)	0.036* (0.022)	0.033 (0.022)	0.032 (0.023)
Observations	3159	3159	3159	3159	3159
Bedridden days in the preceding 2 years	−3.267 (2.225)	−3.118 (2.088)	−3.193 (2.149)	−3.235 (2.290)	−3.179 (2.284)
Observations	3120	3120	3120	3120	3120
Had serious illness in the preceding 2 years	0.021 (0.017)	0.024 (0.017)	0.023 (0.017)	0.020 (0.017)	0.019 (0.018)
Observations	3159	3159	3159	3159	3159
Measured hypertension at interview	−0.040* (0.022)	−0.034 (0.022)	−0.036 (0.022)	−0.046** (0.022)	−0.051** (0.022)
Observations	3099	3099	3099	3099	3099
Panel 2: perceived access to health care					
Get adequate medical care when sick	0.055*** (0.017)	0.053*** (0.017)	0.054*** (0.017)	0.056*** (0.017)	0.057*** (0.017)
Observations	3159	3159	3159	3159	3159
Fail to get necessary care due to costs	−0.030* (0.016)	−0.027* (0.016)	−0.029* (0.016)	−0.031* (0.016)	−0.032* (0.016)
Observations	3158	3158	3158	3158	3158
Panel 3: health spending					
Log(total expenditure +1)	0.069 (0.151)	0.111 (0.147)	0.093 (0.149)	0.045 (0.151)	0.028 (0.151)
Observations	3051	3051	3051	3051	3051
Log(out-of-pocket + 1)	0.082 (0.144)	0.122 (0.144)	0.105 (0.144)	0.058 (0.144)	0.043 (0.142)
Observations	3013	3013	3013	3013	3013

Bootstrap standard errors are reported in parentheses.

*** $p < 0.01$;

** $p < 0.05$;

* $p < 0.1$.

As shown in panel 1, the NCMS has significantly improved ADL and cognitive functions of the elderly enrollees. Specifically, the NCMS take-up status has significantly increased the probability of having no ADL limitation, by 3.8 percentage points, and improved the MMSE scores by 0.932, which is equivalent to an increase of 3.8% over the pretreatment level (24.45). Consistently, the NCMS participants are more likely to have good cognitive function; the estimate is not significant in the main result, but significant at the 10% level in the estimations using bandwidth parameters 0.1 and 0.08. Moreover, the NCMS has also reduced the likelihood of having measured hypertension by 4 percentage points, but only at the 10% significance level. Consistent with Lei and Lin (2009), we find no significant effect of the NCMS on the participants' self-assessed health status.

Panel 2 provides further evidence supporting the beneficial health effects of the NCMS. The results show that the NCMS has significantly improved the participants' perceived access to health care, by increasing the

probability of getting adequate care when sick by 5.5 percentage points and reducing the risk of failing to get necessary care due to costs by 3 percentage points. These results are consistent with the literature (Wagstaff *et al.*, 2009a, 2009b), which also finds a significant positive effect of NCMS on utilization of both outpatient and inpatient care.

In panel 3, there is no evidence to suggest that the NCMS has reduced the enrollees' total medical expenditure and out-of-pocket expenditure. Instead, the estimates are positive but insignificant. Although it may be surprising, this finding is also consistent with the literature (Lei and Lin, 2009; Wagstaff *et al.*, 2009a) and could be explained by the increased healthcare utilization among the NCMS enrollees, as we find in this analysis.

Table IV presents the ATT estimates for different subgroups. In the first three columns, we estimate the impact of the NCMS for different income groups. The results suggest that the poorest 30% (bottom three deciles) benefit the most from NCMS participation in terms of health outcomes and perceived healthcare access. The poor have seen a significant increase in MMSE score and the probability of having good cognitive function and a decrease in the prevalence of hypertension. We also find that the participants from the middle-income groups are less likely to have ADL limitation (at the 10% significance level) and hypertension. The health effects are insignificant for the richest 30%, except for the MMSE score (with marginal significance). Different from the poor, the middle-income and high-income groups have seen no improvement in perceived healthcare access. For each income group, we find no significant effect of NCMS on health expenditure.

The middle two columns show estimates for men and women, respectively. The results suggest that NCMS has improved cognitive functions for both male and female elderly. We also find a significant positive effect of NCMS on ADL status for women, but not for men. However, the main difference between the two groups is that male enrollees are significantly more likely to get adequate medical care and have higher total and out-of-pocket spending, whereas the estimates for women are insignificant.

The last three columns report estimates for China's three regions (eastern, central, and western). Consistent with the findings for the poor elderly, the results show that NCMS has significantly decreased the probability of having hypertension and has improved perceived access to care for the elderly enrollees in the relatively poor western region. Although the elderly in the central region have seen an improvement in perceived access to care, we find no significant health effects of NCMS for them. The evidence is mixed for the elderly enrollees in the relatively affluent eastern region, who have experienced not only an improvement in cognitive functions but also an increase in the prevalence of hypertension and no change of perceived access to care. Consistently, we find no reduction in out-of-pocket spending for each region.

5.2. Effects of NCMS on mortality

As noted earlier, the sample of the elderly has a mortality rate of about 37.5% during the period 2005–2008. This considerable sample loss may result in attrition bias in our main estimates based on the nondeceased sample, if the treated and control groups have different attrition rates due to the direct effect of the NCMS on mortality. Specifically, if the enrolled elderly experienced a lower (higher) mortality rate, then the estimated health benefits of NCMS in our PSMDD estimations are likely to be underestimated (overestimated). Moreover, mortality in itself is an important objective health outcome. Therefore, in this subsection, we further explore the impact of NCMS on mortality of the elderly and discuss the potential sample attrition bias in our main estimation results.

Regarding the analysis of mortality, there are two problems: the first is that we cannot construct a panel data set for those deceased respondents, as their mortality outcomes can only be observed once; the second is that the CLHLS does not ask about NCMS enrollment status of the deceased sample before death. Therefore, we cannot employ the PSMDD strategy here or even assign the deceased sample to treatment or control groups on the basis of individual enrollment status after the implementation of NCMS as we did in the main analysis. Thus, we examine the effect of NCMS on mortality of the elderly in the following two ways.

First, following the literature (Chen *et al.*, 2007; Chang, 2012; Finkelstein *et al.*, 2012; Keng and Sheu, 2013), we employ a difference-in-differences method to estimate the intent-to-treat (ITT) effect of NCMS on

Table IV. PSMDD estimates of NCMS impacts for different population subgroups

	By household income per capita			By gender		By region		
	Poorest 30% (N = 950)	Middle 40% (N = 1226)	Richest 30% (N = 847)	Male (N = 1272)	Female (N = 1757)	Eastern (N = 793)	Central (N = 901)	Western (N = 1335)
Panel 1: health outcomes								
Self-reported health is good	0.013 (0.053)	0.016 (0.048)	-0.044 (0.045)	-0.015 (0.038)	0.002 (0.037)	-0.047 (0.049)	0.020 (0.050)	0.032 (0.050)
No ADL limitation	0.034 (0.037)	0.051* (0.031)	0.032 (0.023)	0.025 (0.026)	0.048** (0.024)	0.039 (0.032)	0.045 (0.031)	0.037 (0.024)
MMSE score	1.523*** (0.556)	0.143 (0.602)	1.187* (0.617)	1.054** (0.492)	0.841* (0.481)	1.373** (0.615)	0.413 (0.661)	1.027 (0.698)
Cognitive function is good	0.069* (0.039)	-0.004 (0.037)	0.046 (0.032)	0.057* (0.034)	0.019 (0.027)	0.079** (0.037)	0.037 (0.034)	0.002 (0.034)
Bedridden days in the preceding 2 years	-3.068 (4.766)	-5.083 (5.460)	-1.560 (2.946)	-5.140 (5.390)	-1.908 (3.662)	-1.867 (3.896)	-3.357 (4.837)	-4.569 (4.915)
Had serious illness in the preceding 2 years	-0.009 (0.041)	0.062 (0.039)	0.002 (0.027)	0.053 (0.033)	0.001 (0.025)	0.038 (0.035)	-0.039 (0.034)	0.036 (0.032)
Measured hypertension at interview	-0.093** (0.040)	-0.104** (0.050)	0.067 (0.044)	-0.032 (0.037)	-0.044 (0.037)	0.094** (0.048)	-0.040 (0.045)	-0.154*** (0.036)
Panel 2: perceived access to health care								
Get adequate medical care when sick	0.127*** (0.036)	0.020 (0.030)	0.022 (0.021)	0.078*** (0.026)	0.038 (0.024)	0.008 (0.023)	0.122*** (0.042)	0.037 (0.025)
Fail to get necessary care due to costs	-0.071** (0.029)	-0.008 (0.021)	-0.015 (0.013)	-0.032 (0.021)	-0.028 (0.017)	-0.029 (0.020)	-0.017 (0.029)	-0.034* (0.019)
Panel 3: health spending								
Log(total expenditure +1)	-0.046 (0.262)	0.066 (0.271)	0.150 (0.206)	0.369* (0.221)	-0.145 (0.170)	0.306 (0.229)	0.452* (0.268)	-0.283 (0.190)
Log(out-of-pocket +1)	-0.132 (0.290)	0.037 (0.276)	0.273 (0.250)	0.428* (0.238)	-0.164 (0.184)	0.390 (0.271)	0.441 (0.357)	-0.303 (0.213)

A bandwidth of 0.06 is used. Bootstrap standard errors are reported in parentheses.

*** $p < 0.01$;** $p < 0.05$;* $p < 0.1$.

3-year mortality, using the full sample (both survivors and the deceased) in waves 2005–2008. The following equation is estimated:

$$Mortality_{it} = \beta_0 + \beta_1 Y2008_t + \beta_2 Target_i + \beta_3 Y2008_t \times Target_i + \beta_4 X_{it} + \varepsilon_{it} \quad (4)$$

where $Y2008_t$ is a binary indicator for wave 2008 and $Target_i$ is an indicator variable for the target population of NCMS, which includes all rural elderly with no pre-NCMS insurance in the experimental counties. As described earlier, in our study sample, no county had the NCMS in 2005, and almost all counties were exposed to it in 2008, when all uninsured rural elderly in NCMS counties were eligible for this program. Those rural elderly who had been covered by any preexisting public health insurance program, such as government employee medical insurance or Urban Employee Basic Medical Insurance,¹² were not subject to this policy change but share the similar environment, so that they provide a natural reference group. We make the comparison regardless of the individual's actual NCMS take-up decision. Therefore, the coefficient β_3 identifies the ITT effect of NCMS on the target elderly population in NCMS counties.

The first part of Table V shows that the 3-year mortality rate of the rural elderly who were uninsured prior to NCMS has been relatively stable during our study period; it was 43.2% in wave 2005 and 42.3% in wave 2008. For those with pre-NCMS public health insurance, the mortality rate increased slightly, from 31.3% in wave 2005 to 34.5% in wave 2008. The simple difference-in-differences estimate of the ITT effect of NCMS is -0.041 , with a standard error of 0.032 , which is not statistically significant. In the last nine columns of Table V, we report the ITT estimates controlling for individual characteristics and region effects identified in Equation 4 for the full sample and different subsamples. Each cell shows the results of a separate regression. These results show that the ITT estimates are negative (except for women and the poor) but insignificant, implying that expanding access to public health insurance has no significant effect on the 3-year mortality rate of the rural elderly. That is consistent with Chen and Jin (2012), who also find that the NCMS has no significant effect on village-level mortality of children and pregnant women.

To further examine the effect of NCMS on the mortality of the elderly who participated in the program, we present the matched posttreatment estimates in the second part of Table V. We use only the study sample of wave 2008¹³ and estimate the impact of NCMS take-up status in 2008 on the probability of dying in the next 3 years. The propensity score matching (PSM) results show that the NCMS reduces the 3-year mortality rate for the participants by 3.7 percentage points, which is significant at the 10% level. The negative impact of NCMS enrollment is stronger in eastern regions. Consistent with the finding in Taiwan (Keng and Sheu, 2013), we also find a negative effect of NCMS on mortality for women (significant at the 10% level), but not for men.

Although the PSM estimation helps control only for observable differences, not unobservable differences between the treatment and control groups, we have to admit that the matched posttreatment estimates may suffer from the bias due to unobservable factors. However, taken together, these findings in Table V help rule out the possibility that the NCMS has shortened the life of the seriously ill elderly who participated in it. Thus, although we exclude the deceased respondents in the panel data analyses in Section 5.1, it would not lead to upward-biased estimates of NCMS impacts on self-reported and measured health outcomes.

5.3. Testing the assumption of PSMD estimation

To identify the effects of the NCMS in our main analysis in Section 5.1, we rely on the PSMD method to control for observables, time-constant unobservables, and time-varying unobservables of common trend. Although we have obtained high-quality matching based on a comprehensive set of variables and conducted difference-in-differences estimations on the matched sample, there are still concerns that our estimates may

¹²The Urban Employee Basic Medical Insurance is employment based and mainly covers employees and retirees in both the public and private sectors, regardless of residence and *Hukou*.

¹³Although we know whether the elderly respondents in wave 2005 died or not in the next 3 years, we do not know their NCMS enrollment status before death in the CLHLS. So we cannot use the study sample of wave 2005 here.

Table V. Estimates of NCMS impact on three-year mortality

Wave	3-year mortality rate	Sample size	All	Poorest 30%	Middle 40%	Richest 30%	Male	Female	Eastern	Central	Western
Part A. DID estimation using waves 2005–2008: intent-to-treat effect of NCMS for the full sample and different subsamples											
Treated: no pre-NCMS insurance	0.432	[5805]									
Control: has government employee insurance or urban employee basic insurance	0.423	[7113]	–0.005	0.092	–0.001	–0.016	–0.026	0.037	–0.001	–0.016	–0.007
Observations	0.313	[320]	(0.029)	(0.088)	(0.058)	(0.037)	(0.037)	(0.046)	(0.036)	(0.069)	(0.078)
R^2	0.345	[812]									
Part B. PSM estimation using wave 2008: average treat effect of NCMS for the full sample and different subsamples											
Treated: enrollees	0.407	[5029]	–0.037*	–0.055	–0.011	–0.051	–0.021	–0.041*	–0.094***	–0.030	–0.018
Control: unenrolled in NCMS	0.461	[2084]	(0.021)	(0.036)	(0.024)	(0.036)	(0.028)	(0.024)	(0.034)	(0.033)	(0.027)
Observations		[7113]	6153	1629	2661	1863	2630	3523	2453	1613	2087

Robust standard error are reported in parentheses in Part A. Bootstrap standard errors are reported in parentheses in Part B. Other covariates include gender, age, age squared, years of schooling, marital status, an indicator of having a white-collar job before age 60, an indicator of co-residing with adult children, number of children alive, household income per capita, and regional dummies.

*** $p < 0.01$;

** $p < 0.05$;

* $p < 0.1$.

Table VI. Test of the assumption of PSMDD estimation using 2002–2005 panel data

	Bandwidth				
	0.06	0.1	0.08	0.04	0.02
Panel 1: health outcomes					
Self-reported health is good	−0.032 (0.037)	−0.047 (0.037)	−0.040 (0.036)	−0.021 (0.037)	−0.009 (0.037)
Observations	1607	1607	1607	1607	1607
No ADL limitation	0.003 (0.017)	−0.001 (0.017)	0.001 (0.017)	0.004 (0.017)	0.002 (0.017)
Observations	1645	1645	1645	1645	1645
MMSE score	0.581 (0.375)	0.578 (0.386)	0.590 (0.382)	0.549 (0.366)	0.513 (0.363)
Observations	1645	1645	1645	1645	1645
Cognitive function is good	0.012 (0.030)	0.013 (0.030)	0.013 (0.030)	0.008 (0.029)	0.005 (0.030)
Observations	1645	1645	1645	1645	1645
Bedridden days in the preceding 2 years	2.641 (2.864)	3.080 (2.895)	2.900 (2.856)	2.263 (2.985)	1.687 (3.327)
Observations	1634	1634	1634	1634	1634
Had serious illness in the preceding 2 years	−0.029 (0.025)	−0.027 (0.024)	−0.028 (0.025)	−0.031 (0.025)	−0.034 (0.025)
Observations	1645	1645	1645	1645	1645
Measured hypertension at interview	0.016 (0.047)	0.012 (0.048)	0.014 (0.048)	0.019 (0.047)	0.025 (0.048)
Observations	1621	1621	1621	1621	1621
Panel 2: perceived access to health care					
Get adequate medical care when sick	−0.012 (0.021)	−0.012 (0.022)	−0.012 (0.021)	−0.012 (0.021)	−0.010 (0.021)
Observations	1645	1645	1645	1645	1645
Fail to get necessary care due to costs	0.034* (0.018)	0.036* (0.020)	0.035* (0.020)	0.031* (0.017)	0.030* (0.017)
Observations	1644	1644	1644	1644	1644
Panel 3: health spending					
Log(total expenditure + 1)	—	—	—	—	—
Observations	—	—	—	—	—
Log(out-of-pocket + 1)	—	—	—	—	—
Observations	—	—	—	—	—

Bootstrap standard errors are reported in parentheses. The treatment status is defined on the basis of waves 2005–2008. The 2002 wave did not ask respondents about their medical expenditure.

*** $p < 0.01$;

** $p < 0.05$;

* $p < 0.1$.

be driven by different unobserved, time-varying heterogeneities between the treatment and control groups. To conduct a falsification test, we use the 2002–2005 data of the same treatment and control groups in the main analysis¹⁴ and define the treatment/control status according to their NCMS enrollment in wave 2008 as before. This restriction leads to a balanced panel subsample of 1748 respondents (2002–2005). The treatment group consists of 1231 respondents, and the control group of 517 respondents. Because neither the treatment nor the control group was exposed to NCMS during 2002–2005, we expect no significant differences between the two groups in health outcomes and healthcare access from PSMDD estimations.¹⁵

¹⁴Refer to Appendix Table AI in the online supporting information for the detailed description of sample construction for this falsification test.

¹⁵The 2002 wave did not ask respondents about their health spending. So we cannot obtain the estimates for health spending during 2002–2005.

Table VI reports the analogous estimates of the effects of the (nonexistent) NCMS using the 2002–2005 data. As clearly shown, the ATT estimates either are insignificant or have the wrong signs for all dependent variables. This specification test thus falsifies the possibility that some unobserved time-variant heterogeneity drives the impacts of NCMS established in Section 5.1.

6. CONCLUSIONS

In this paper, we explore the effectiveness of NCMS in improving health outcomes and reducing health spending for the elderly in rural China, using panel data drawn from CLHLS 2005–2008. We employ the PSMDD strategy to identify the effects of NCMS coverage. We find that NCMS enrollment has had a significant positive effect on ADL and cognitive function of the rural elderly. These results are robust to a variety of specifications. Although the literature finds insignificant health effects of the NCMS for the sample of the general population (Lei and Lin, 2009; Chen and Jin, 2012), it is possible that the NCMS is more effective for the rural elderly population, who are most susceptible to health risk (Levy and Meltzer, 2008) and have more elastic demand for health care (Ringel *et al.*, 2002). Furthermore, although the NCMS has no significant effect on 3-year mortality for the previously uninsured elderly in experimental counties, there is moderate evidence that it is associated with reduced mortality for the elderly enrollees, especially those in eastern regions. It is also found that the effects of the NCMS on elderly health are differentiated by socioeconomic status and by gender. In particular, the NCMS has a larger impact on health outcomes and perceived access to care for the elderly at the bottom of the income distribution, suggesting that the NCMS also helps reduce gaps in access and health among the rural elderly.

To provide further insights into the health benefits of the NCMS, we also show that the NCMS has significantly improved perceived access to health care for the elderly enrollees. Although we cannot distinguish between outpatient care use and inpatient care use, these results are consistent with previous findings that the NCMS has significantly increased the enrollees' utilization of outpatient services and inpatient services (Wagstaff *et al.*, 2009a, 2009b; Yu *et al.*, 2010) as well as preventive care services (Lei and Lin, 2009). Moreover, consistent with Lei and Lin (2009) and Wagstaff *et al.* (2009a), we still find no significant effects of the NCMS on total healthcare expenditure and out-of-pocket expenditure in most specifications, probably because of constrained funding and low reimbursement rate (Yi *et al.*, 2009).

It must be acknowledged that this study has several potential limitations. First, although the combination of matching and difference-in-differences methods helps control for observables and unobservables, there may still be a potential confounding influence of unobserved time-varying factors. Although we have to admit this limitation of our method, the falsification test provides some support for its validity and shows that this potential bias may not be a problem in this study. Second, owing to lack of information on the NCMS enrollment status of the deceased sample before death, we cannot adopt a pre–post treatment–control study design for mortality outcome and estimate the effect of the NCMS for the enrollees (i.e., the ATT). Although we have conducted an ITT analysis for mortality outcome and found that the estimated ITTs of the NCMS are negative and insignificant, the effects of the NCMS on mortality may be stronger in the ATT estimations. The posttreatment matched difference results show a modest negative effect of the NCMS on elderly enrollees' mortality. Given that, the impacts of the NCMS on other health outcomes might be underestimated. Third, we do not have detailed, objective measurements of spending for outpatient and inpatient services. The CLHLS data for health expenditures in the previous year were self-reported and relatively crude and may be subject to measurement errors. Moreover, the reimbursement rate of the NCMS has been increasing over time since 2009 (Chen and Jin, 2012). We leave it to future research to explore fully whether and how the NCMS impacts financial risk, with the availability of micro-data on detailed health expenses.

Overall, despite its relatively short life and limited financing, our study provides empirical evidence for the beneficial health effects of the NCMS, which may have important policy implications for the deepening healthcare system reform in China, and also for the development of health insurance in other developing countries. It motivates further studies to investigate the health effects of the NCMS among general populations and other vulnerable populations and to better understand the underlying mechanisms.

CONFLICT OF INTEREST

The authors have no conflicts of interest to declare.

ACKNOWLEDGEMENTS

This study was supported by NIH R01 grant (5R01-AG023627) and grants from the Natural Science Foundation of China (no. 71110107025 and no. 71203244), the National Social Science Fund of China (no. 13CJY028), Beijing Higher Education Young Elite Teacher Project (no. YETP0989), and Program for New Century Excellent Talents in University (no. NCET-12-0993). We would like to thank Xiaoyan Lei, Yaohui Zhao, Zhibiao Liu, and seminar participants in the National School of Development and Peking University for their helpful comments. All errors are our own.

REFERENCES

- Babiarz KS, Miller G, Yi H, Zhang L, Rozelle S. 2010. New evidence on the impact of China's new rural cooperative medical scheme and its implications for rural primary healthcare: multivariate difference-in-difference analysis. *British Medical Journal* **c5617**: 341–350.
- Babiarz KS, Miller G, Yi H, Zhang L, Rozelle S. 2012. China's new cooperative medical scheme improved finances of township health centers but not the number of patients served. *Health Affairs* **31**(5): 1065–1074.
- Beckett S, Gould W, Lillard L, Welch F. 1988. The panel study of income dynamics after fourteen years: an evaluation. *Journal of Labor Economics* **6**(4): 472–492.
- Brown PH, de Brauw A, Du Y. 2009. Understanding variation in the design of China's new cooperative medical system. *The China Quarterly* **198**: 304–329.
- Card D, Dobkin C, Maestas N. 2008. The impact of nearly universal insurance coverage on health care utilization and health: evidence from Medicare. *American Economic Review* **98**(5): 2242–2258.
- Chang S. 2012. The effect of Taiwan's National Health Insurance on mortality of the elderly: revisited. *Health Economics* **21**(11): 1257–1270.
- Chen Y, Jin GZ. 2012. Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. *Journal of Health Economics* **31**(1): 1–14.
- Chen L, Yip W, Chang MC, Lin HS, Lee SD, Chiu YL, Lin YH. 2007. The effects of Taiwan's national health insurance on access and health status of the elderly. *Health Economics* **16**(3): 223–242.
- Chou SY, Grossman M, Liu JT. 2011. The impact of national health insurance on birth outcomes: a natural experiment in Taiwan. *Working Paper* 16811, National Bureau Economic Research, Cambridge, MA.
- Crum RM, Anthony JC, Bassett SS, Folstein MF. 1993. Population-based norms for the Mini-Mental State Examination by age and educational level. *The Journal of the American Medical Association* **269**(18): 2386–2391.
- Currie J, Gruber J. 1996a. Health insurance eligibility, utilization of medical care and child health. *Quarterly Journal of Economics* **111**(2): 431–466.
- Currie J, Gruber J. 1996b. Saving babies: the efficacy and cost of recent changes in the Medicaid eligibility of pregnant women. *Journal of Political Economy* **104**(6): 1263–1296.
- Currie J, Gruber J. 1997. The technology of birth: health insurance, medical interventions and infant health. *Working Paper* 5985, National Bureau Economic Research, Cambridge, MA.
- Currie J, Decker S, Lin W. 2008. Has public health insurance for older children reduced disparities in access to care and health outcomes? *Journal of Health Economics* **27**(6): 1567–1581.
- Decker SL, Remler DK. 2004. How much might universal health insurance reduce socioeconomic disparities in health? A comparison of the US and Canada. *Applied Health Economics and Health Policy* **3**(4): 205–216.
- Ding W, Lehrer SF. 2010. Estimating treatment effects from contaminated multiperiod education experiments: the dynamic impacts of class size reductions. *The Review of Economics and Statistics* **92**(1): 31–42.
- Du L, Zhang W. 2007. The Development on China's Health, No. 3. Social Academic Press: Beijing China.
- Eggleston K, Ling L, Meng Q, Lindelow M, Wagstaff A. 2008. Health service delivery in China: a literature review. *Health Economics* **17**(2): 149–165.
- Ekman B. 2004. Community-based health insurance in low-income countries: a systematic review of the evidence. *Health Policy and Planning* **19**(5): 249–270.

- Feng X, Tang S, Bloom G, Segall M, Gu X. 1995. Cooperative medical schemes in contemporary rural China. *Social Science & Medicine* **41**(8): 1111–1118.
- Finkelstein A, McKnight R. 2008. What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. *Journal of Public Economics* **92**(7): 1644–1668.
- Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, Allen H, Baicker K, Oregon Health Study Group. 2012. The Oregon health insurance experiment: evidence from the first year. *The Quarterly Journal of Economics* **127**(3): 1057–1106.
- Folstein MF, Folstein SE, McHugh PR. 1975. “Mini-mental state”: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research* **12**(3): 189–198.
- Franks P, Clancy C, Gold MR, Nutting PA. 1993. Health insurance and subjective health status: data from the 1987 National Medical Expenditure Survey. *American Journal of Public Health* **83**(9): 1295–1299.
- Frölich M. 2004. Finite-sample properties of propensity-score matching and weighting estimators. *Review of Economics and Statistics* **86**(1): 77–90.
- Gakidou E, Lozano R, Gonzalez-Pier E, Abbott-Klafter J, Barofsky JT, Bryson-Cahn C, Feehan DM, Lee DK, Hernandez-Llamas H, Murray CJ. 2006. Assessing the effect of the 2001–06 Mexican health reform: an interim report card. *The Lancet* **368**(9550): 1920–1935.
- Gu D. 2008. General data quality assessment of the CLHLS. In *Healthy Longevity in China, The Springer Series on Demographic Methods and Population Analysis*, vol. **20**, Yi Z, Poston DL, Vlosky DA, Gu D (eds.). Springer: Dordrecht, the Netherlands.
- Gu D, Zeng Y. 2004. Data assessment of the Chinese Longitudinal Healthy Longevity Survey (in Chinese). In *Determinants of Healthy Longevity of the Oldest-old in China*, Zeng Y, Xiao Z, Zhang C (eds.). Peking University Press: Beijing, China.
- Gu D, Dupre ME, Liu G. 2007. Characteristics of the institutionalized and community-residing oldest-old in China. *Social Science & Medicine* **64**(4): 871–883.
- Hadley J. 2003. Sicker and poorer—the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. *Medical Care Research and Review* **60**(2 supp): 3S–75S.
- Hadley J, Waidmann T. 2006. Health insurance and health at age 65: implications for medical care spending on new Medicare beneficiaries. *Health Services Research* **41**(2): 429–451.
- Heckman J, Ichimura H, Todd P. 1998. Matching as an econometric evaluation estimator. *Review of Economic Studies* **65**(223): 261–294.
- Heckman J, LaLonde R, Smith J. 1999. The economics and econometrics of active labor market programs. In *Handbook of Labor Economics*, Ashenfelter O, Card D (eds.). Elsevier: Amsterdam; 1865–2097.
- Imbens G. 2004. Nonparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics* **86**(1): 4–29.
- Jowett M, Deolalikar A, Martinsson P. 2004. Health insurance and treatment seeking behavior: evidence from a low-income country. *Health Economics* **13**(9): 845–857.
- Katz S, Ford AB, Moskowitz RW, Jackson BA, Jaffe MW. 1963. Studies of illness in the aged: the index of ADL: a standardized measure of biological and psychosocial function. *The Journal of the American Medical Association* **185**(12): 914–919.
- Keng S, Sheu S. 2013. The effect of national health insurance on mortality and the SES–health gradient: evidence from the elderly in Taiwan. *Health Economics* **22**(1): 52–72.
- Lei X, Lin W. 2009. The new cooperative medical scheme in rural China: does more coverage mean more service and better health? *Health Economics* **18**(S2): S25–S46.
- Leuven E, Sianesi B. 2012. Psmatch2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. *Statistical Software Components*. Available from: <http://ideas.repec.org/c/boc/bocode/s432001.html> [Accessed 08/26/12].
- Levy H, Meltzer D. 2008. The impact of health insurance on health. *Annual Review of Public Health* **29**: 399–409.
- Liu Y, Hsiao WCL, Li Q, Liu X, Ren M. 1995. Transformation of China’s rural health care financing. *Social Science & Medicine* **41**(8): 1085–1093.
- McWilliams JM, Zaslavsky AM, Meara E, Ayanian JZ. 2004. Health insurance coverage and mortality among the near-elderly. *Health Affairs* **23**(4): 223–233.
- Mihelic AH, Crimmins EM. 1997. Loss to follow-up in a sample of Americans 70 years of age and older: the LSOA 1984–1990. *The Journals of Gerontology: Psychological Sciences and Social Sciences* **52B**(1): S37–S48.
- Ministry of Health. 2004. Research on National Health Services: An Analysis Report of National Health Services Survey in 2003 (in Chinese). Xie He Medical University Press: Beijing.
- National Bureau of Statistics. 2011. Statistical Communiqué of the People’s Republic of China on the 2010 National Economic and Social Development (in Chinese).

- Panopoulou G, Velez C. 2001. Subsidized health insurance, proxy means testing and the demand for health care among the poor in Colombia. Colombia Poverty Report Volume II. Washington DC, World Bank.
- Ringel JS, Hosek SD, Vollaard BA, Mahnovski S. 2002. The elasticity of demand for health care; a review of the literature and its application to the military health system. Available from: http://www.rand.org/pubs/monograph_reports/MR1355.html [Accessed 12/26/11]
- Rosenbaum P, Rubin D. 1983. The central role of the propensity score in observational studies for causal effects. *Biometrika* **70**(1): 41–55.
- Rosenbaum P, Rubin D. 1985. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician* **39**(1): 33–38.
- Sepehri A, Sarma S, Simpson W. 2006. Does non-profit health insurance reduce financial burden? Evidence from the Vietnam Living Standards Survey Panel. *Health Economics* **15**(6): 603–616.
- State Council. 2002. Decisions of the State Council on strengthening rural healthcare (in Chinese). State Council, China.
- Sun X, Jackson S, Carmichael GA, Sleigh AC. 2009. Catastrophic medical payment and financial protection in rural China: evidence from the new cooperative medical scheme in Shandong province. *Health Economics* **18**(1): 103–119.
- Sun X, Sleigh AC, Carmichael GA, Jackson S. 2010. Health payment-induced poverty under China's new cooperative medical scheme in rural Shandong. *Health Policy and Planning* **25**(5): 419–426.
- Tang S, Meng Q, Chen L, Bekedam H, Evans T, Whitehead M. 2008. Tackling the challenges to health equity in China. *The Lancet* **372**(9648): 1493–1501.
- Trujillo AJ, Portillo JE, Vernon JA. 2005. The impact of subsidized health insurance for the poor: evaluating the Colombian experience using propensity score matching. *International Journal of Health Care Finance and Economics* **5**(3): 211–239.
- Wagstaff A, Lindelow M. 2008. Can insurance increase financial risk? The curious case of health insurance in China. *Journal of Health Economics* **27**(4): 990–1005.
- Wagstaff A, Yu S. 2007. Do health sector reforms have their intended impacts? The World Bank's Health VIII project in Gansu province, China. *Journal of Health Economics* **26**(3): 505–535.
- Wagstaff A, Lindelow M, Gao J, Xu L, Qian J. 2009a. Extending health insurance to the rural population: an impact evaluation of China's new cooperative medical scheme. *Journal of Health Economics* **28**(1): 1–19.
- Wagstaff A, Yip W, Lindelow M, Hsiao W. 2009b. China's health system and its reform: a review of recent studies. *Health Economics* **18**(S2): S7–S23.
- Wang H, Yip W, Zhang L, Hsiao WC. 2009. The impact of rural mutual health care on health status: evaluation of a social experiment in rural China. *Health Economics* **18**(S2): S65–S82.
- Yi H, Zhang L, Singer K, Rozelle S, Atlas S. 2009. Health insurance and catastrophic illness: a report on the new cooperative medical system in rural China. *Health Economics* **18**(S2): S119–S127.
- Yip W, Hsiao WC. 2009. Non-evidence-based policy: how effective is China's new cooperative medical scheme in reducing medical impoverishment? *Social Science & Medicine* **68**(2): 201–209.
- You X, Kobayashi Y. 2009. The new cooperative medical scheme in China. *Health Policy* **91**(1): 1–9.
- Yu B, Meng Q, Collins C, Tolhurst R, Tang S, Yan F, Bogg L, Liu X. 2010. How does the new cooperative medical scheme influence health service utilization? A study in two provinces in rural China. *BMC Health Services Research* **10**(1): 116.
- Zeng Y, Vaupel JW, Xiao Z, Zhang C, Liu Y. 2001. The healthy longevity survey and the active life expectancy of the oldest old in China. *Population: An English Selection* **13**(1): 95–116.
- Zeng Y, Vaupel JW, Xiao Z, Zhang C, Liu Y. 2002. Sociodemographic and health profiles of the oldest old in China. *Population and Development Review* **28**(2): 251–273.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publishers web site.