NBER WORKING PAPER SERIES

THE LARGEST INSURANCE EXPANSION IN HISTORY:
SAVING ONE MILLION LIVES PER YEAR IN CHINA

Jonathan Gruber
Mengyun Lin
Junjian Yi

Working Paper 31423
http://www.nber.org/papers/w31423

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
July 2023

We are grateful to Ben Olken and participants at Boston University, Harvard, and MIT Health Economics Seminar and Chinese University of Hong Kong Economics Seminar for helpful comments. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Largest Insurance Expansion in History: Saving One Million Lives Per Year in China
Jonathan Gruber, Mengyun Lin, and Junjian Yi
NBER Working Paper No. 31423
July 2023
JEL No. H4,I13

ABSTRACT

The New Cooperative Medical Scheme (NCMS) rolled out in China from 2003-2008 provided insurance to 800 million rural Chinese. We combine aggregate mortality data with individual survey data, and identify the impact of the NCMS from program rollout and heterogeneity across areas in their rural share. We find that there was a significant decline in aggregate mortality, with the program saving more than one million lives per year at its peak, and explaining 78% of the entire increase in life expectancy in China over this period. We confirm these mortality effects using micro-data on mortality, other health outcomes, and utilization.

Jonathan Gruber
Department of Economics, E52-434
MIT
77 Massachusetts Avenue
Cambridge, MA 02139
and NBER
gruberj@mit.edu

Mengyun Lin
The Wang Yanan Institute for
Studies in Economics
Xiamen University
lin.mengyun@xmu.edu.cn

Junjian Yi
National School of Development
Peking University
junjian@nsd.pku.edu.cn

An online appendix is available at http://www.nber.org/data-appendix/w31423

1 Introduction

Medical spending risk is already one of the greatest sources of financial risk in developed countries of the world. Fortunately, virtually every developed country has either government or privately-provided insurance programs that cover the vast majority of the population. In 2018, across all OECD nations, as a result, the share of healthcare costs borne out of pocket amounts to only 20.1% of healthcare costs, and 1.7% of incomes.[1]

In developing nations, health insurance coverage of the population is more sporadic. Traditionally, this has not been a priority for development because both healthcare quality and expenditure were low. In 2018 for example, across all low-income countries (LIC), 44.1% of healthcare expenditure was out of pocket, and this amounted to 3.45% of incomes.[2]

The need for health insurance has changed dramatically in the developing world in recent decades as higher quality medical techniques have been extended worldwide. The benefit of improved medical access in the developing world is one key source of dramatic improvements in life expectancy around the world. Over the past two centuries, every region of the world has seen its life expectancy double; the global average life expectancy today is higher than any country in 1950.[3] Even the least developed countries have seen their average life expectancy grow by almost 25 years since 1960.[4] But this improved medical care comes at a cost. Total medical expenditures in low-income countries have risen from 4.2% to 5.3% of GDP over the past 20 years.[5]

[1] 20.1%: https://stats.oecd.org/Index.aspx?ThemeTreeId=9. This is an unweighted average across all OECD countries. 1.7%: https://stats.oecd.org/Index.aspx?DataSetCode=AV_AN_WAGE. This is an unweighted average across all OECD countries. We take per capita out-of-pocket (OOP) expenditure and divide by the average annual wage in each country and average these percentages.

[2] 44.1%: https://data.worldbank.org/indicator/SH.XPD.OOPC.CH.ZS?locations=XM. This is an average of OOP expenditure as a percentage of healthcare expenditure of low-income countries as designated by the world bank.

3.45%: https://data.worldbank.org/indicator/NY.ADJ.NNTY.PC.CD?end=2016&locations=XM &start=2016&view=bar. This is an unweighted average across all low-income countries as designated by the world bank (Eritrea, North Korea, Somalia, South Sudan, Syria, and Yemen are removed due to incomplete information). We take per capita OOP expenditure and divide by adjusted net national income per capita in each country and average these percentages.

[3] https://ourworldindata.org/life-expectancy-globally#:~:text=Globally%20the%20life%20expectancy %20increased,more%20than%20twice%20as%20long.

[4] https://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=XL.

[5] https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?locations=XM.

As medical costs rise, there has been a dramatic expansion of public insurance programs in developing countries. Table 1 lists the introduction of new insurance programs in some of the largest developing countries over the past 30 years, along with estimates of the number of individuals enrolled in these programs. The largest single expansion in Table 1 is the New Cooperative Medical Scheme (NCMS) program that was initiated in China in 2003. This program provided comprehensive coverage of hospital expenditures for rural households in China. Within five years of its introduction, it was covering over 800 million persons, making it the largest insurance program in the history of the modern world. Expenditure on the program was 0.43% of GDP, or 7.16% of healthcare expenditure, in 2015.[6]

Despite its rank as the largest insurance expansion in modern history, there is little evidence on the impacts of the NCMS on health outcomes. There are only a few articles in both economics and public health literature assessing the impact of the NCMS on health outcomes (e.g., Cheng et al. 2015; Zhou et al., 2017). These papers generally find little benefit for health from this enormous program, consistent with limited positive findings for a number of massive insurance expansions in the developed world. But, as we discuss below, these papers suffer from a number of concerns over identification and precision. In particular, there is strong evidence of endogenous timing of program adoption, as counties that adopted the programs at different times were on different health trajectories, which biases the estimates of program effects.

We provide a comprehensive evaluation of the impact of the NCMS on the health of rural residents in China. We combine multiple sources of data to paint a broad picture of the effect of NCMS on mortality and other health outcomes, as well as on medical care utilization and expenditure. We rely on the fact that this program was rolled out at different times across areas, and that its impact across areas is proportional to the share of the eligible agricultural population. The share of counties covered by the program rose from 11% in 2003 to 95% by 2008. Meanwhile, across counties in China, there is wide variation in the agricultural population eligible for this program; for example, in the sample of counties that we use for

[6] Data on NCMS expenditure is from China Health Statistical Yearbook, 2009 & 2016; data on China's GDP is from China Statistical Yearbook, 2009 & 2016.

our mortality analysis, the agricultural share varies from 5% to 96%. By using both of these dimensions of variation, we are able to control for endogenous adoption of the program. We depend on a triple-difference framework to identify the policy effects and the identification assumption is that, in the absence of the policy, the difference in health outcomes between higher and lower agricultural share counties should evolve in parallel regardless of adoption time.

We first use aggregate data on mortality rate and life expectancy at the county level to obtain an intent-to-treat estimate of health effects of NCMS. We find that the passage of the law created enormous benefits for the health of the rural Chinese population. Within eight years of passage by 2010, the program had enrolled 88% of China's rural population and 62% of the country's total population.[7] We estimate that, by 2010, the NCMS reduced the mortality rates by 12.0% (95% CI=[2.2%, 22.7%]) and increased the life expectancy by 2.7% (95% CI=[0.2%, 5.2%]). This led directly to a reduction in the mortality rate among the enrolled population by 19.3% (12%/62%), and extended life expectancies by 4.4% (2.7%/62%), or 78% of the 2.5-year increase in life expectancies from 2003-2010.[8] Since 2008, we estimate this program saved more than one million lives per year, a population larger than seven states in the US. The estimated cost per life saved of 66.1 thousand RMB ($9.8 thousand) is well below both estimates of the value of a statistical life in China and other measures such as compensation for accidental death.[9]

We then explore the mechanisms through which this program achieved its success, using two micro-data sets on healthcare utilization and health outcomes among large samples of the Chinese population. This analysis allows us to estimate treatment-on-the-treated estimates of those who actually enrolled in the program, and to examine a rich array of health outcomes, healthcare utilization, and protection from out-of-pocket (OOP) expenditure.

Micro-data estimates show that the program had even larger effects on elderly mortality than estimated in our aggregate mortality data. We also find that the program dramatically

[7] China Statistical Yearbook 2011 and China Health Statistical Yearbook 2003-2011

[8] The change in life expectancy from 2003 to 2010 is from China Statistical Yearbook 2011.

[9] The monetary values here are in 2010 prices. The exchange rate in 2010: 1 USD = 6.77 RMB. The same rule is applied to all monetary values throughout the paper, unless stated otherwise.

improved healthcare utilization; the probability of utilizing medical care increased by up to 77%. Measures of health status such as limitations of activities of daily living, cognitive deficits, and self/interviewer-assessed health improved dramatically as well, confirming that the life extension we find reflects broader improvements in health—and are consistent with the role of the hospital as the nexus of care in the modern Chinese health care system.

We then explore three mechanisms through which these results may have arisen. We show that OOP expenditure fell substantially for those at the top of the medical expenditure distribution, providing protection against catastrophic expenditures. We show that, at most, a small share of the effect is arising through endogenous supply responses.

We explain the estimated large health effects of NCMS in the context of the Chinese healthcare system. In particular, the comprehensive inpatient coverage provided by the NCMS plays a critical role for general health care, not just the most acute cases, in China. First of all, inpatient admission does not follow the type of detailed referral system in the US—rather, individuals are typically sent for inpatient admission from outpatient primary care clinics (Yip and Hsiao, 2008). Indeed, 90% of hospital visits come directly from outpatient referrals (Milcent,2018).[10] Second, hospitals are widely used for management of mild disease and chronic care; for example, one-quarter of hospitalizations in China are for chronic and not acute illnesses, compared to about one-tenth in the US.[11] Third, there is almost no charity care or uncompensated care provided by Chinese hospitals.[12] Without an NCMS card, individuals must pay upfront, or they will not be admitted, even to the emergency department. If their initial deposit is close to being exceeded, the patients must "top up" their account or face discharge (Milcent, 2018). For these reasons, having a new form of insurance through NCMS could have a broad spectrum of effects on health.

Our paper proceeds as follows. Section 2 provides background on health transformation in the developing world. Section 3 describes the NCMS program and its place in the trajectory

[10]2020 China Health Statistical Yearbook.

[11]Chronic diseases include diabetes, hypertension, liver cirrhosis, cholecystography, and rheumatoid arthritis. Data for China is from 2018 China Health Statistical Yearbook; data for the US is from McDermitt & Roemer (2021).

[12]See evidence in Section 7.1.

of healthcare development in China. Section 4 describes our multiple sources of data and our empirical strategy. Section 5 carries out our aggregate analysis of mortality rates. Section 6 uses micro-data to explore impacts on healthcare utilization and self-reported health. Section 7 then turns to the mechanisms behind our results, while Section 8 concludes.

2 Background on Health Transformation in China and Other Developing Nations

China's healthcare system has gone through a radical transformation since the founding of the People's Republic of China in 1949 (Burns and Liu, 2017). At that point, healthcare provision in China was dominated by traditional medicine, and life expectancy was low. In 1952, total healthcare expenditure was only 1.3% of GDP, compared to 5% in the US in 1960.[13] Life expectancy in China was 35 in 1949, and it was even lower in rural areas.[14]

In the 70 years since 1949, the Chinese healthcare system has undergone a fundamental shift towards Western medicine (Milcent, 2018). As a result, healthcare expenditure has risen to 6.6% of GDP in 2018.[15] And China has seen one of the most amazing improvements in life expectancy in the modern world, with an increase in life expectancy to 76 by 2015 (Feng, 2013).[16]

But as China and other countries have modernized their healthcare systems, they have faced a fundamental question of how to share this burden within society (Yip and Hsiao, 2008). Some countries, such as India, Pakistan, and Singapore, have decided to place the burden mostly on households, with the share of expenditure coming out of pocket above 50% in these countries. Other countries, such as Thailand and South Africa, have moved to a model comparable to many European nations, with only about 10% of costs borne out of pocket. Nations such as Vietnam, China, Peru, and Brazil have followed a middle ground, with roughly 25-33% of healthcare expenditure being out of pocket.

[13]Research Report on China Healthcare Expenditure, Ministry of Health 2008 (in Chinese).
[14]China Health Statistical Yearbook 2019.
[15]China Health Statistical Yearbook 2019.
[16]China Statistical Yearbook 2016.

These developing nations have undertaken a variety of approaches to covering the burden of healthcare expenditure. One approach is to reduce the user fees often attached to the use of public medical services. Gertler et al. (1987), Litvack and Bodart (1993), and Souteyrand et al. (2008) explore user fees via financing proposals in Peru, a field experiment in Cameroon, and an HIV/AIDS treatment in developing countries, respectively; all find equity, quality and access concerns are reduced by the elimination of fees. Gertler et al. (1987), in particular, finds that the loss in consumer welfare mainly concentrates on the poor. Zombré et al. (2017) find that removing user fees for young children in Burkina Faso increased usage dramatically.

Most comparably for our study, there is a very large literature evaluating large-scale insurance expansions in developing nations (see Appendix 1). Taken together, this large literature suggests that enormous insurance expansions in developing countries can play a major role in reducing OOP expenditure and in increasing utilization. But evidence on health impacts is surprisingly limited. More generally, there are remarkably few comprehensive evaluations of the impact of health insurance reform on health outcomes, healthcare utilization, and OOP expenditure.

Given its size, there is a surprisingly modest literature evaluating the NCMS on health. We are aware of only a few economics papers that focus primarily on the health impacts of the program, and their identification is mainly based on the staggered adoption of NCMS across counties. For example, Lei and Lin (2009) use CHNS data and find little impacts on self-reported health status using fixed-effect model, instrumental variable approach, and difference-in-differences with propensity score matching method; Chen and Jin (2012) use a cross-sectional approach in aggregate data and find little effect of NCMS enrollment on mortality; Cheng et al. (2015) use difference-in-differences estimator with propensity score matching, and find positive effects NCMS on health outcomes, such as activities of daily living, and a small reduction in mortality outcomes; Huang and Liu (2023) use a difference-in-differences framework and find that exposure to the NCMS during ages 0-5 significantly improves health and cognitive outcomes during adolescence.

The closest paper to ours in the public health literature is Zhou et al. (2017), who carry out a difference-in-differences study on aggregate mortality data using the timing of NCMS expansion across rural counties. They find a decrease in mortality rates, but the causal link is weak. As we explain below, the substantial variation in exposure of counties to the program, as well as underlying trends in non-rural counties, explain the much stronger effects we find.

There is a larger literature on other effects of the program. Several studies find increases in healthcare utilization: Long et al. (2012) find a higher use of Cesarean-section delivery, and Wagstaff et al. (2009) find higher inpatient and outpatient utilization but a limited decrease in OOP expenditure, while Sun et al. (2009) find a sizable reduction in catastrophic medical expenditure.[17] Liu (2016) suggests that NCMS helps households to mitigate adverse outcomes associated with health shocks in household finance; Bai and Wu (2014) find that NCMS increases household consumption by 5%; Yip and Hsiao (2009) find that NCMS plays a modest role in reducing medical impoverishment.

3 The New Cooperative Medical Scheme

NCMS is the successor of the Cooperative Medical Scheme (CMS). The CMS was established in the 1950s alongside the establishment of the People's Republic of China. The CMS covered up to more than 90% of rural residents from the early 1950s to the economic reform in 1979 (Burns and Liu, 2017). However, the CMS began to collapse after 1979 when the rural communes were dissolved under market-oriented economic reform. As a result, the CMS coverage of rural residents declined sharply from nearly universal to almost zero in the 1990s (Blumenthal and Hsiao, 2015).

In the absence of any health insurance and qualified medical providers in rural areas, individuals or families became the only payer for all medical costs in the 1990s (Milcent, 2018). The share of OOP expenditure paid by individuals and families out of the total healthcare expenditure at the national level tripled, from 20% in 1979 to nearly 60% in 2000; the 30-fold rise in OOP expenditure from 1970 to 2000 outpaced the 4.5-fold rise in income.[18] With the

[17] This is consistent with our findings that OOP expenditure decreases substantially for those at the top of total expenditure distribution.

[18] China Health Statistical Yearbook 2001 and China Statistical Yearbook 2001.

rising OOP expenditure, rural families were fully exposed to health risks and felt a great financial burden, which put pressure on the Chinese government to implement a new public health insurance program for rural families (Burns and Liu, 2017).

In 2003, China initiated the NCMS program. This new program was administered separately in each of around 2,800 county-level administrative units in China since most rural residents are scattered across the townships and villages governed by the county government.[19] Under broad guidelines issued by the central government, at least two to three pilot counties were selected in each province in the first year, and more counties were gradually included, aiming at achieving full coverage by 2010.[20] The rollout of the program across counties is shown in Figure 1. In 2003, the first year when the program was available, only 11% of counties implemented it. There followed a rapid expansion so that by 2008, coverage was virtually universal in all counties. Coverage started to fall after 2008 because some counties replaced the NCMS with a new scheme covering both rural and urban residents after a parallel program started to roll out in urban areas (at a much smaller scale)(Yip, 2020).

In practice, as we show in Figure A1 in Appendix 2, both the agricultural share and the average GDP were generally higher among early-adopting counties than late-adopting counties. This implied that the program was first adopted by highly rural counties with relatively strong economies, then was extended to less strong economies and ultimately to less rural counties. This complicates analysis if these different types of counties are on different time trends, which we address below.

The NCMS targeted rural residents, and agricultural Hukou serves as the eligibility requirement. The Hukou system is a household registration system implemented after the establishment of the People's Republic of China. This system categorizes each Chinese citizen into two types: an agricultural Hukou holder or a non-agricultural Hukou holder. The Hukou is used to determine one's eligibility for social service and welfare to his registered place of residence. One's Hukou type is inherited from parents, and its change is highly restricted

[19]On average, there are 14 townships in one county, and each township contains 16 villages.

[20]The official document (in Chinese) is archived at http://www.gov.cn/zwgk/2005-08/12/content_21850.htm.

by the government. According to the China Population Census, the share of agricultural Hukou remained relatively flat during the first ten years of the 2000s, at 72.31% in 2000 and 70.86% in 2010; due to the rapid industrialization of China over the past decade, it has fallen to 55.62% today.[21] This rigorous assignment system implies that, during our study period, migration responses do not drive any of the results that we show below.[22]

The central government issued broad guidelines for the design and implementation of the NCMS program; meanwhile, local governments in each county had considerable discretion over details such as premiums and coverage of expenditure. As a result, the program's benefits packages vary geographically. Generally, the program focuses on inpatient care, the largest source of healthcare spending risk, and provides a "catastrophic" insurance coverage design (Yip and Hsiao, 2008). During the initial years, individuals who were hospitalized were responsible for a deductible of around 200-500 RMB (about $25-62), which was about 6-15% of net annual income per capita in 2005 (Burns and Liu, 2017). Beyond this deductible, there was a coinsurance rate of 30-50%. And the overall reimbursement is capped at 15,000 RMB (about $1,875), roughly five times net annual income per capita of rural residents in 2005 (Bai and Wu, 2014). As time went by, the NCMS benefits became more generous, with the coinsurance rate falling to 10-30% (Burns and Liu, 2017).

In addition, the program provided more heterogenous coverage of outpatient care. In particular, 18% of counties provided payment for outpatient care in 2006; in another 65% of counties, the program set up mandatory household medical savings accounts that would be used to pay for outpatient care (Burns and Liu, 2017).[23] The accounts are funded by the individual payment for NCMS premiums, which started as roughly 10 RMB and rose to roughly 20 RMB by 2012 (Burns and Liu, 2017). Over time, this savings mechanism has been gradually abolished and replaced by a more typical outpatient reimbursement program, with

[21] http://www.stats.gov.cn/english/PressRelease/202002/t20200228_1728917.html.

[22] Migrants rarely use healthcare services in the destination county because across-city/province healthcare claims are not reimbursable.

[23] Establishing and contributing to the household savings account was mandatory in counties that applied this scheme. The amount of contribution, which was by reallocation of the total premiums collected from both individuals and the government, was determined by local governments, and varied across counties (Milcent, 2018).

a reimbursement rate of 50-80% (Milcent, 2018). Therefore, the vast majority of counties introduced both inpatient and outpatient coverage of some type.

The NCMS was paid for by an insurance premium. The central government set a minimum level, and the local government adjusted it according to local income and benefits packages. At the start of the program, the minimum premium was 30 RMB ($3.63) per year, of which the government (the individual) paid 2/3 (1/3). By 2008, the minimum premium had risen to 100 RMB ($14.50), of which the government paid 80%. In counties with savings accounts for outpatient expenditures, individual premiums were the funding source for the savings accounts.

Participation in this program was voluntary. However, due to the modest premiums, intense government subsidies, and strong government mobilization ability, participation rates were as high as 75% in the first three years after initiation and increased to 96% in 2010. Individuals who paid the premium and enrolled in the program received a certificate card that could be used when visiting clinics and hospitals.

There was a significant debate over the potential efficacy of the NCMS. Concerns were raised along three dimensions. The first was that the benefits were not generous enough. The incomplete coverage of outpatient care and the relatively high cost sharing would mitigate any health benefits (Yip and Hsiao, 2008). The second was that participation was voluntary, leading to low enrollment and potential adverse selection (Milcent, 2018). The third was that the decentralized administration of the NCMS at the county level would be problematic because risk pools would not be large enough and local governments would not have the capacity to support the program (Burns and Liu, 2017). For this reason, the health benefits of NCMS were an open question.

In addition, health insurance coverage was extended through a new program introduced for urban areas in 2008, Urban Resident Basic Medical Insurance (URBMI). This program was much smaller than the NCMS, targeting 210 million unemployed urban residents in 2010, which is only one-fourth of those covered by the NCMS. In 2009, the government started integrating the URBMI with the NCMS to form a new public health insurance program, Urban and Rural Residents Basic Medical Insurance (URRBMI) (Milcent, 2018).

4 Data and Empirical Strategy

4.1 Data

We use a variety of data sets for our analysis. To define our treatment variable, we focus on China's 2000 Population Census. This provides the share of the population with agricultural Hukou in each county. The distribution of this share across counties in China is shown in Figure 2.

For aggregate mortality, we turn to the China Death Surveillance Point Dataset (DSP). This is a nationally representative data set which records all deaths and population counts by sex and age. The data includes information from a sample of counties throughout China. Round 1, from 1991-2000, included 145 counties, while Round 2, from 2004-2012, included 161 counties. We hand collected data on the dates of NCMS implementation, and were able to gather that data for 121 of the 161 counties included in Round 2.[24]

Our key dependent variables are age-adjusted mortality rates and life expectancy at birth at the county/year level.[25] We focus our analysis on the years 2004 (the first year available in the second panel) to 2010. A potential confounder of our analysis is the new insurance program introduced for urban areas in 2008 which, while much smaller, weakens the comparison to some extent with non-rural counties. As we show below, excluding the later years strengthens our results.

We also collect data on healthcare supply at the county level, including healthcare staff and hospital beds. We use these data to examine the potential mechanism of supply responses to the NCMS. In addition, we control for local economic development, such as

[24]We actively searched through official announcements of NCMS implementation from local governments' official websites and news reports. We compared the characteristics of counties with missing information on the starting year of NCMS and those of the remaining counties using China's 2000 Population Census. Table A1 in Appendix 3 shows no significant difference between the two groups.

[25]We obtain the county-level mortality rates by age from recorded deaths and population counts. These mortality rates are used to calculate the life expectancy at birth and age-adjusted mortality rates. The age adjustment is based on direct standardization (Curtin and Klein, 1995), where the standardized population by age is from China's 2000 Population Census.

GDP per capita and the average wage of urban employees, and a variety of other government policy interventions that may be coincident with NCMS, such as social aid centers and local investment in fixed assets. These data are from China Statistical Yearbooks (county-level) from 2005 to 2011.

The means and standard errors from these data are shown in Table A2 in Appendix 4. For comparability, we also show means for these variables for the nation as a whole as well. We find that, along every dimension, there are no significant differences between our sample and the entire nation.

We then turn to micro-data analyses of individual impacts. For this, we use two data sets. The Chinese Longitudinal Healthy Longevity Survey (CLHLS) provides data from individual face-to-face interviews in seven waves during 1998-2014 from 22 provinces in China.[26] Survivors in each wave are re-interviewed, while the deceased are replaced by new participants to maintain sample balance; however, family members of the deceased are interviewed for information about the deceased, such as the date of death.[27] The sample began with those elderly above age 80, then added those aged 65-79 since 2002, and included the elders' adult children aged 35-64 in the 2002 and 2005 waves. For our analysis, we include individuals aged 65-110 for all waves from 1998-2014, except for new participants in 2014;[28] our final sample for the mortality analysis includes 78,446 observations of 33,307 unique individuals, among which 19,857 died as of 2014 at an average age of 96.

We use several key variables from these data, as summarized in the first panel of Table A3 in Appendix 4. The survey asks whether the individual suffered from a serious illness in the last two years, whether they have limitations in their activities of daily living,[29] their mental

[26]The survey was conducted in 1998, 2000, 2002, 2005, 2008, 2011, and 2014, respectively. The official website of the CLHLS: https://cpha.duke.edu/research/chinese-longitudinal-healthy-longevity-survey-clhls.

[27]The CLHLS also provides information on the cause of death, but it is not available for waves since 2008.

[28]Sample age is restricted because data on mortality and health status before dying between waves were only collected for the elders aged 65-110.

[29]Limited ADL is an indicator that equals one if the individual reports a need of help with eating, dressing, moving, toileting, bathing, and continence.

health status,[30] their self-reported health status and corresponding interviewer-reported health status.[31] Questions regarding healthcare utilization include whether the individuals consider themselves as getting adequate medical services, whether they refuse medical services due to lack of money, and their total medical expenditure during the previous year.

The second is the China Health and Nutrition Survey (CHNS). This is a longitudinal survey of about 7,200 households, including over 30,000 individuals in 15 provinces that vary substantially in geography, economic development, public resources, and health indicators.[32] The survey uses a multistage, random cluster process to draw the sample of households, which ensures its national representativeness.[33] There are ten waves from 1989 to 2015. We use waves 2000, 2004, 2006, 2009 and 2011 in our analysis, and exclude the six provinces that are not covered by the survey until 2011. Our final sample consists of 19,037 individuals residing in 225 communities of 54 counties in nine provinces.[34]

Each wave of the CHNS includes three modules for individual, household, and community. The modules include a variety of measures summarized in the second panel of Table A3 in Appendix 4. The survey provides information about activities of daily living [35] and

[30] We describe one's mental health status based on questions about the frequency of the following five emotions: (1) looking on the bright side of things; (2) being happy as younger; (3) feeling fearful and anxious; (4) feeling lonely or isolated; and (5) feeling useless with age. Options include "Always", "Usually", "Sometimes", "Seldom", and "Never". Mental health is an indicator that "Always" or "Usually" is reported for the first two positive emotions, and "Seldom" or "Never" is reported for the last three negative emotions.

[31] Both self-reported and interviewer-reported health status are categorical variables but on different scales. Self-reported health status is rated on a 1-5 scale: 1 for "Very good", 2 for "Good", 3 for "So-so", 4 for "Bad", and 5 for "Very bad". Interviewer-reported health status is on a 1-4 scale: 1 for "Surprisingly healthy", 2 for "Relatively healthy", 3 for "Moderately ill", and 4 for "Very ill". We define self-reported health as an indicator for "Very good" or "Good" health, and interviewer-reported health as an indicator for "Surprisingly healthy" or "Relatively healthy". The correlation between these two measures is 0.378.

[32] The provinces covered in the survey changed across waves. There were nine provinces from 2000 to 2009. Three mega cities (at the same administrative level as province) were included in 2011, and another three provinces in 2015.

[33] The official website of the CHNS: https://www.cpc.unc.edu/projects/china.

[34] In China, there are five administrative levels: central government, province, prefecture, county, and community. Community, which refers to villages, townships, or neighborhoods in counties, is the lowest-level administrative unit.

[35] Limited ADL is an indicator that equals one if the individual has any limitation in bathing, dressing, eating, toileting, combing hair, shopping, cooking, using public transportation, managing money, and using telephone.

cognitive ability[36] for those above 55 in waves 2000 to 2006. In addition, the survey records information on healthcare utilization during the past four weeks, including any use of medical services when feeling sick or getting injured, any inpatient or outpatient care, and any use of preventive care. The survey also asks about the total medical expenditure and the proportion paid by health insurance. This allows us to compute the OOP expenditure and the ratio of OOP expenditure to total expenditure. Finally, the CHNS has questions on whether the individuals consume alcohol or cigarettes, and the amount of consumption per day.

4.2 Empirical Strategy

A natural starting point for evaluating the phase-in of a program like NCMS is a difference-in-differences (DD) framework as in Zhou et al. (2017), using variation in implementation dates across counties. But this program has a major disadvantage: the timing of program adoption was not random. In particular, more agricultural and higher-income counties were the first to adopt the program, and it is likely that these counties were on unparallel trajectories before the program began.

We illustrate this point in two ways in Appendix 5. First, we show that potentially omitted variables are strongly correlated with the implementation timing of NCMS across counties in the DD estimation approach. Table A5 shows the estimation results from a typical staggered DD framework.[37] We begin by showing that the NCMS has no effects on age-adjusted mortality rate and life expectancy in columns (1) and (2). We then replicate this regression with four other dependent variables and their corresponding lags: county GDP, county average wage of urban employees, investment in fixed assets per capita, and the number of social aids centers per 10,000 people. In every case, we find a highly significant

[36]Questions on cognitive ability differ across waves. We choose six questions that are constantly asked in all three waves, including one counting exercise and five subtraction exercises. We define an indicator of cognitive deficit that equals one if the individual did not answer all the questions correctly.

[37]In particular, we estimate the following specification

$$Y_{ct} = \alpha_1 Post_{ct} + \delta_c + \delta_t + \epsilon_{ct},$$

where Y_{ct} refers to the outcome variables; $Post_{ct}$ is an indicator of county c implementing NCMS in year t; δ_c and δ_t are county fixed effects and year fixed effects, respectively.

correlation between these variables or their lagged values and whether the county has adopted the NCMS, suggesting that other factors may be correlated with the adoption of the reform (Table A6).[38]

Second, we find that there are unparallel pre-trends that do not satisfy the identification assumption in the DD estimation. In particular, we classify counties that adopted NCMS before 2006 as the early NCMS adopters, and other counties as the late NCMS adopters. We then use the first round of DSP data from 1991 to 2000 to show the pre-trends for the two groups in Appendix 5 Figure A2.

Due to these limitations, we adopt a richer empirical framework that allows for endogenous timing of program adoption. The key innovation is recognizing that there is an additional source of variation not exploited by previous analysis: the agricultural density of the county. The NCMS program will have larger effects in counties with more agricultural density. This allows us to compare counties that adopted at the same time, but who have different levels of agricultural density.[39]

In particular, for the aggregate DSP data, we estimate an equation of the form:

$$\log(Y_{cpt}) = \alpha_0+\alpha_1 AgriShare_{c,2000}\times Post_{ct}+AgriShare_{c,2000}\times\delta_t+\mathbf{X}'_{\mathbf{ct}}\boldsymbol{\alpha}+\delta_c+\delta_{pt}+\delta_{t_c^0,t}+\delta_c\times t+\epsilon_{cpt}, \quad (1)$$

where Y_{cpt} denotes age-adjusted mortality rate (the number of deaths per 100,000) or the life expectancy at birth for county c in province p in year t; $AgriShare_{c,2000}$ is the share of the population with an agricultural Hukou in county c in the year 2000; $Post_{ct}$ indicates that county c has adopted NCMS in year t; and $\mathbf{X_{ct}}$ includes GDP per capita and the average wage of urban employees, and an interaction of agricultural share with a set of year fixed

[38]We do the same analysis with our preferred specification in Equation (1), and results in Table A7 show no significant coefficients on the county-level variables or their lags.

[39]In the same spirit, we incorporate in the micro data analysis the comparison of residents with and without an agricultural Hukou as an additional source of identification.

effects.[40] Standard errors are clustered at the county level, and regressions are weighted by county population.

We include a rich set of fixed effects. We include county-fixed effects (δ_c) to capture any fixed differences across counties, province-by-year fixed effects (δ_{pt}) to capture general time patterns at the province level, as well as an interaction of agricultural share with a set of year fixed effects ($AtriShare_{c,2000} \times \delta_t$) to capture varying time patterns associated with the different agricultural shares.[41]

As noted above, the counties that adopted NCMS at different times are different along key dimensions; if those differences are fixed, they will be captured by the county fixed effects; if those differences are changing over time (in a manner correlated with our dependent variables) but the time patterns are common for counties in the same province or of the same agricultural share, they will be captured by the province-by-year fixed effects or agricultural share interacted with year dummies; otherwise, these time-variant unobserved confounders could bias our estimates.

We account for this in two ways. First, we allow each county to have its own linear time trend ($\delta_c \times t$). Second, we include a separate set of year fixed effects for each NCMS timing group (i.e., NCMS-timing-by-year FE, $\delta_{t_c^0,t}$), which allows for nonlinear changes over time among the counties that adopt the program earlier or later.[42] As we also control for the interaction of agricultural share with a set of year fixed effects, we compare the health outcomes of counties with high agricultural shares to those with low agricultural shares

[40]We use a log transformation of our dependent variables primarily due to the presence of outliers. For instance, the mean mortality rate in our sample is 546 (deaths per 100,000 population), with values ranging from a minimum of 55 to a maximum of 1,130. The log transformation helps to mitigate the influence of outliers on our estimations, leading to more stable and precise estimates. Our baseline findings remain generally robust when using the raw levels instead of the logs of the dependent variables. Specifically, the estimates from regressions for the age-adjusted mortality rate are negative and marginally statistically significant, with t-statistics exceeding 1.5; the estimates from regressions for life expectancy are positive and statistically significant. These results are in Appendix 6 Table A10.

[41]In particular, the interaction of agricultural share with a set of year fixed effects helps to control for confounding contemporary reforms targeted rural areas and agricultural population, such as the elimination of agricultural tax since 2006. To address the concern regarding potentially insufficient variations in the agricultural share within each adoption cohort, we examine the summary statistics of the agricultural share by the year of adoption, as illustrated in Table A4 in Appendix 4. Our findings indicate substantial variation within each cohort, with the agricultural share ranging from 0.1 to 0.9.

[42]NCMS-timing-by-year fixed effects, in other words, indicate that we have separate year effects for all counties that implemented in 2004, 2005, etc.

that simultaneously adopted the NCMS, while controlling for any time-varying differences in health outcomes solely associated with agricultural shares.

Compared to the conventional DD framework, the identification assumption of our empirical strategy is weaker. For illustration purposes, we justify our identification assumption by discretizing $AgriShare_{c,2000}$. Then, Equation (1) is similar to a triple-difference estimation, where we measure the relative health outcomes between the counties with high versus low agricultural shares that simultaneously adopted the NCMS, and compare the change of the relative outcomes among counties that adopted NCMS at different times. The identification assumption is thus weaker than that for a standard DD estimation (Gruber, 1994). In particular, the latter requires that the timing of NCMS adoption is exogenous; but the former allows for the endogenous timing, as the variable of $Post_{ct}$ is directly controlled for.[43] In particular, the triple-difference strategy requires no contemporaneous shocks that differently affect the evolvement of relative outcomes of high versus low agricultural share counties by the time of NCMS adoption. We rigorously verify this identification assumption in Appendix 5.

Our identification assumption is that the health outcome gaps between high and low agricultural share counties would have evolved similarly in the absence of reform, regardless of the time when NCMS was implemented. However, this assumption may be violated due to time-varying factors. These factors may include the existing and evolving rural-urban gaps in health outcomes and healthcare facilities, and local government policies that are designed to mitigate these gaps. Such factors can differentially impact the timing of NCMS adoption in counties with high or low agricultural shares (as indicated by $HighAgri_{c,2000} \times Post_{ct}$). Consequently, the omission of county-level, time-varying unobservable factors may introduce bias into the triple-D estimator. This argument is corroborated by Appendix 5 Figure A3, using DSP data from 1991 to 2000.

Our empirical strategy seeks to address with this issue by controlling for NCMS-timing-by-year FE, county-specific linear trend, and province-by-year FE. The crux is whether the

[43] We control for $Post_{ct}$ by controlling for $\delta_{t_c^0,t}$, NCMS-timing-by-year FE.

omitted unobservables could be appropriately proxied by these fixed effects and trends. This is an empirical question. Figure A3 in Appendix 5 demonstrates that the trends for both mortality rate and life expectancy are indeed parallel when we remove the county-specific linear trend and fixed effects within the NCMS adoption cohort and province.

5 Aggregate Impacts

5.1 Aggregate Analysis using DSP

Our initial results are presented graphically in Figure 3, which shows log mortality rate and log life expectancy over time. These graphs are in event study format, where year zero is the year that a county adopted NCMS, and the sample is divided into counties with above or below agricultural shares at the median. We first regress log mortality rate and log life expectancy on NCMS-timing-by-year fixed effects and a county-specific linear trend. We then plot the residuals from this regression relative to the year of adoption; observations more than 3 (5) years before (after) the NCMS adoption are binned into groups.

The results are striking. Both groups of counties have similar (relatively flat) trends before the program, with the more urban counties having a significantly lower mortality rate (by about 8%) and a longer life expectancy (by about 1.5%).

But these differences quickly shrink as soon as the NCMS is put into place, and then reverse. While mortality is slightly rising and life expectancy is slightly falling in more urban counties after the year of NCMS adoption, mortality is substantially reduced, and life expectancy is dramatically extended in more rural counties. By the time a new steady state is achieved 2-3 years after passage, mortality rates are roughly 4% lower in more rural counties, and life expectancy is roughly 0.7% higher in more rural areas.

The regression estimates corresponding to Equation (1) are shown in the first column of Table 2. We see a substantial drop in mortality associated with counties with a high agricultural share post the NCMS, relative to before. We estimate that living in a county with the median agricultural share (78.5%), relative to a county with no agricultural population, implied a decline in mortality of 14.0% (95% CI=[2.6%, 25.5%]) at the time of the NCMS.[44] Similarly, we show a strong and significant correlation between agricultural share and the change in life expectancy after the NCMS. Living in a county with the median agricultural share resulted in a rise in life expectancy of 3.2% (95% CI=[0.21%, 6.1%]) relative to a county with a zero agricultural share.[45]

We also estimate the dynamic version of this specification:

$$\log(Y_{cpt}) = \beta_0 + \sum_{j=-3, j\neq -1}^{j=5} \beta_j 1\{t - t_c^0 = j\} \times AgriShare_{c,2000} + AgriShare_{c,2000} \times \delta_t + \mathbf{X}'_{\mathbf{ct}}\boldsymbol{\beta} + \delta_c + \delta_{pt} + \delta_{t_c^0,t} + \delta_c \times t + \epsilon_{cpt}, \tag{2}$$

where t_c^0 is the first year when county c adopted the NCMS. The dummy variable $1\{t-t_c^0 = j\}$ indicates that t is j years relative to t_c^0. The year before NCMS adoption is omitted, so the estimates are normalized to zero in that year. Other variables are defined in Equation (1). The results are shown in Figure 4. We see that there is no effect before the passage of the NCMS, and a gradually growing effect thereafter, consistent with our causal interpretation.[46]

5.2 Robustness

We highlighted above the lack of robustness of the DD approach to omitted variables and pre-trends. In this section, we show that our approach is much more robust to these concerns and other issues. We show these results in the remaining columns of Table 2 and Appendix 6.

One concern is that adoption of NCMS was part of a suite of actions by local governments to improve outcomes during this period of tremendous growth for China. To address this concern, we include as controls the number of social aid centers and the investment in fixed assets per capita; column (2) shows that this has little impact on our results. In column (3), we move the other direction by removing the set of covariates (GDP per capita and average

[44] $14.0\% = (1 - exp(-0.197)) \times 0.785$, and below is the same.

[45] We also examine the effect of NCMS on two most frequently observed causes of death: respiratory disease and cerebral-cardiovascular diseases. Results in Appendix 6 Table A11 show that both drop significantly in counties with a higher agricultural share post the NCMS than before.

[46] As one moves away from the event, the confidence intervals become wider. Possibly, this is due to the panel being unbalanced and having fewer observations at the ends.

wage) described earlier; the results are nearly identical, confirming the independence of our identifying variation from other factors.[47]

There is a possibility that the improvement of health outcomes associated with relatively high agricultural share reflects the influence of other county characteristics that are also correlated with NCMS timing and the evolution of health outcomes. To rule out this alternative, we additionally control for the interactions of Post NCMS with county variables measured in the year 2000, including GDP, urban wages, fixed asset investment, and the number of social aid centers, respectively. The results in the first four columns of Appendix 6 Table A12 do not deviate from our baseline, indicating that agricultural share has a unique impact.

We also examine the confounding influence of other contemporary reforms. The first is Province-Manage-County reform that enhanced the fiscal capacity of local government since the early 2000s. We include local government fiscal income as additional control variable. The second is trade internalization as a result of China's entry into the WTO in 2001. We control for the interaction of year dummies with prefecture tariff gap following Erten and Leight (2021). The results are in the last two columns of Appendix 6 Table A12, suggesting that our baseline findings are not confounded by these contemporary reforms.

In column (4) of Table 2, we shorten the sample period to 2004-2007 to remove any influence of the 2008 introduction of an urban insurance program; our results get stronger but less precise when we do so, and we cannot reject that they are the same as for the full period.[48]

We also consider the possibility that our findings are driven by outliers in three ways. First,

[47] We further examine the robustness of our results to omitted variable bias, using the method described by Oster (2019). We assume that R^2, including both unobserved and observed variables, is 1 because the 1.3 times real R2 exceeds 1 in our case. We then compute the ratio of selection on unobserved variables to selection on observed variables that match a treatment effect of zero. We find the Oster ratios for both control variables (GDP per capita and average wage) are negative, suggesting that the estimated treatment effect is unlikely to be driven by unobserved variables.

[48] The process of data collection for the DSP data underwent a change in 2008. Prior to this, the mortality data was collected by the local CDC based on the county's registered (hukou) population; following this change, the data was reported to a national information system and was based on the county's residential population. It should be noted that the residential population of a county is highly correlated with its registered population. By using data prior to 2008, we are able to alleviate concerns that the shift in data collection might confound the analysis.

in column (5), we drop the four counties in the top and bottom 1% of our residual distribution; this does not change our findings. Second, we use bootstrapping by randomly sampling counties from our data with replacement for 1,000 times and re-estimating Equation (1). Figure A5 in Appendix 6 shows the distribution of the estimated coefficients on $Post \times AgriShare2000$. The means and standard errors of the estimates do not deviate from the baseline, and 90% are negative (positive) in the case of mortality rates (life expectancy). Third, we implement a permutation test by randomly reassigning the adoption dates to counties and re-estimating Equation (1) for 1,000 times. Figure A6 in Appendix 6 shows that fewer than 5 percent of placebo replications produce significant effects ($p < 0.05$) for either outcome. The permutation test indicates $p < 0.05$ for the null hypothesis that there is no treatment effect on either mortality rate or life expectancy.

Finally, we use the earlier mortality panel that runs from 1994 to 2000 to conduct a placebo test on our results. We take each implementation data for NCMS and shift it back by 10 years, so that the program is (counterfactually) phased in from 1994 to 1998. We then redo our analysis on this earlier data. As shown in the last column of Table 2, we find no effect on either of our dependent variables. As an alternative, we randomly reassign the counterfactual adoption time and re-estimate the regression for 1,000 times. The distribution of estimates in Appendix 6 Figure A7 also shows no effects in over 98% of placebo scenarios.

Recent research has highlighted the sensitivity of estimation results to the way unit-specific linear trends are calculated. Specifically, the trends may be inadvertently capturing some of the treatment effects of the program as they are estimated over the entire period (Neumark et al., 2014; Wolfers, 2006). To check the robustness of our findings when controlling for county-specific time trends, we employ alternative methods described by Agha and Zeltzer (2022) and Goodman-Bacon (2021). Fortunately, we have access to two rounds of DSP data, with the first round covering the period from 1991 to 2000. Thus, we can use the first round to impute alternative measures for county-specific time trends. Our analysis reveals that our main results remain robust when controlling for these alternative measures, as shown in Table A14 in Appendix 6.

Although the framework in Equations (1) and (2) are widely used, recent literature shows that interpretation of such estimates is complicated in the presence of treatment effect heterogeneity.[49] To alleviate the concern, we apply the estimator based on De Chaisemartin and D'Haultfoeuille (2022a) and the estimation yields an average effect of -0.272 ($s.e. = 0.067$) on mortality rate, and 0.039 ($s.e. = 0.017$) on life expectancy as of year 2007 for counties that adopted the NCMS after 2005.[50] Such sample restriction is based on the proposed estimation method. To establish the comparability between the new estimates and our baseline estimates, we re-estimate Equation (1) using the same restricted sample. The estimate of β_j is -0.285 ($s.e. = 0.136$) for mortality rate, and 0.047 ($s.e. = 0.023$) for life expectancy, suggesting that our baseline estimates are quite robust to heterogeneous treatment effect.

Furthermore, we estimate a triple-difference version of the estimator proposed by De Chaisemartin and D'Haultfoeuille (2022a) and Sun and Abraham (2021) in Appendix 6.[51] Although estimation results in Figure A8 and Table A13 are not directly comparable to our baseline estimates, they help to validate the common trend assumption and the causal effects of NCMS adoption on health outcomes.

5.3 Interpretation

To estimate the total impact of the NCMS, we use the specification shown in the first column of Table 2. For each year, we estimate the result based on the average agricultural share of the counties that have adopted the NCMS. Doing so, we estimate that over the 2004-2007 period, the NCMS saved 550,000 lives per year, and from 2008 onwards was saving more

[49]See a review in De Chaisemartin and D'Haultfoeuille (2022b).

[50]To the best of our knowledge about the recent trending DID literature, we are not aware of an estimation method that could estimate dynamic treatment effects with continuous treatment, as in Equation (2). de Chaisemartin and D'Haultfoeuille (2022a) proposes an estimator for DID with continuous treatment, but it is intended to estimate only the average treatment effect.

[51]To utilize the estimator in De Chaisemartin and D'Haultfoeuille (2022a), we first obtain the residuals by regressing the health outcomes on covariates and fixed effects; we then compute the differences between the average residuals of the high and the low agricultural counties that adopted the NCMS at the same time and use them as dependent variables in the estimation using the STATA command (*did_multiplegt*). When employing the estimator proposed by Sun and Abraham (2021), we estimate a cohort-specific dynamic treatment effects interacted with the indicator of high agricultural share. The assumption underlying these two estimators aligns with that of our empirical specification—in the absence of reform, the health outcome gaps between high and low agricultural share counties within each cohort would have followed parallel trajectories.

than one million lives per year.[52]

Similarly, based on our estimates the NCMS raised the life expectancy of the entire nation of China by 1.94 years from 2003 to 2010.[53] Over this period, the life expectancy of China rose by 2.5 years.[54] This suggests that the implementation of the NCMS can explain 78% of the entire rise in life expectancy over this period.

We can also assess the cost effectiveness of NCMS based on our data. Expenditures on NCMS in 2008 were 63.7 billion (in 2010 RMB, hereafter the same), and we estimate that 1.012 million lives were saved that year for the whole nation. That implies a cost of 63.0 thousand RMB ($9,300) per life saved, completely ignoring any of the additional medical benefits that we show below.[55]

How does this compare to typical estimates of the value of a life? Prior studies show that the value of a statistical life in rural China varies from 148-296 thousand RMB (Hammit and Zhou, 2006) to 600 thousand-1.0 million RMB (Wang and He, 2014). The legal compensation for death from injury in China is set to 20 times the local dispensable income per capita, which amounts to 122 thousand RMB in rural China in 2008.[56] The limit on compensation for death by a traffic accident in liability insurance is 110 thousand RMB.[57] By any of these metrics, this program was highly cost-effective.

[52]We estimate that the NCMS reduced the mortality rate by 2.7%-11.6% (95% CIs=[0.5%,4.9%]-[2.1%,21.0%])between 2004 and 2007, and 12.0% (95% CIs=[2.2%,21.7%]) from 2008 onwards. We then calculate the number of lives saved per year during these two periods with the baseline mortality rate of 582 per 100,000 in 2003.

[53]We estimate that the NCMS extended life expectancy by 2.7% (95% CI=[0.2%, 5.2%]) by 2010, which was 1.94 years based on the life expectancy of 72.4 in 2003.

[54]We use life expectancies from the China's 2000 and 2010 Population Census and 2005 1% Sampling Census, and compute the life expectancies in other years between 2000 and 2010 by linear projection.

[55]Using the same method, we compute the average cost per life saved between 2004 and 2010 is 66.1 thousand RMB ($9.8 thousand).

[56]This standard of compensation is set by the Supreme Court in the Interpretation of the Supreme People's Court on Several Issues of the Applicable Law in Hearing Cases of Compensation for Personal Injury (in force since 2004). The per capita disposable income rural households in 2008 is 6.1 thousand according to China Statistical Yearbook 2009.

[57]Regulations on Compulsory Insurance for Traffic Accident Liability of Motor Vehicles.

6 Effects on Utilization and Other Health Outcomes

In this section, we dive into more details on the impacts of the NCMS on healthcare utilization and on other measures of health outcomes. This allows us to confirm that our mortality outcomes reflect increased utilization of healthcare resources and broader improvements in underlying health.

6.1 CLHLS Results

The 1998-2014 waves of the CLHLS allow us actually measure NCMS enrollment as an outcome variable, so that we can estimate treatment-on-the-treated IV models, where we instrument NCMS enrollment by an interaction term between NCMS passage in that county and individuals' agricultural Hukou status. Since the Hukou type is not available in CLHLS, we follow the literature and assign an agricultural Hukou to an individual if he or she lives in rural areas and has no access to government-sponsored welfare schemes for residents with urban Hukou (Huang and Zhang, 2021).[58]

We begin by confirming our mortality results using the microdata. In particular, we estimate a mortality hazard model of the form:

$$h_c(m|S_{icm}) = h_{0,c}(m)exp(\beta_1 NCMS_{icm} + \beta_2 NCMS_{cm} + \beta_3 AgriHukou_i + \mathbf{X_i'}\boldsymbol{\beta} + \delta_{t_c^0,t_m}), \quad (3)$$

where $h_c(m|S_{icm})$ is the hazard of death for individual i with characteristics S_{icm} in county c in year m relative to the first year when individual i participated in the survey; $h_{0,c}(m)$ is the baseline hazard of death, which is assumed to be constant in a given county to account for all unobserved county-specific health-related heterogeneities; S_{icm} includes an indicator for NCMS enrollment ($NCMS_{icm}$), an indicator for NCMS county ($NCMS_{cm}$), a

[58]The CLHLS does not ask about the deceased's NCMS enrollment status; we assume their NCMS status remains the same as it was in the last wave when they were alive. It is possible that this imputed status may introduce measurement error. According to China's 2000 Population Census, the correlation between agricultural Hukou status and residing in rural areas is as high as 0.99. We further investigate the NCMS enrollment rate of individuals without rural hukou in our imputation. It turns out that less than 3% of them were enrolled in NCMS in the CLHLS sample.

dummy for having an agricultural Hukou ($AgriHukou_i$), individual characteristics ($\mathbf{X_i}$), and NCMS-timing-by-year fixed effects ($\delta_{t_c^0,t_m}$). Individual characteristics include dummies for age, gender, minority status, household size, housing income per capita, drinking, smoking, and an indicator of getting adequate medical services when in serious illness in childhood. All these characteristics are held constant at their baseline values.

The NCMS enrollment status might be endogenous because participating in the program is voluntary. For example, health endowments and risk preferences, which are unobserved by researchers, may differ between enrollees and non-enrollees. To address this concern, we use the county NCMS adoption interacted with one's agricultural Hukou status as an instrument for the individual's NCMS enrollment status, $NCMS_{icm}$, in Equation (3). This instrument is valid because it utilizes the staggered adoption of NCMS across counties and a pre-determined eligibility criterion of Hukou status. The 2SLS estimates, however, are inconsistent because Equation (3) is nonlinear. Instead, we use the control function (CF) method (Wooldridge, 2015).[59] Similar to the 2SLS method, the CF method has two stages. The first-stage specification is

$$NCMS_{icm} = \alpha_0+\alpha_1 NCMS_{cm}\times AgriHukou_i+\alpha_2 NCMS_{cm}+\alpha_3 AgriHukou_i+\mathbf{X_i'}\boldsymbol{\alpha}+\delta_c+\delta_{t_c^0,t_m}+\varepsilon_{icm}, \tag{4}$$

where $NCMS_{icm}$ indicates that individual i in county c has enrolled in the NCMS in year m relative to the first year when the individual i participated in the survey; $NCMS_{cm}$ indicates that county c has adopted the NCMS in year m, and $AgriHukou_i$ indicates that individual i has an agricultural Hukou. We control for the same set of controls as in Equation (3) along with county fixed effects (δ_c).[60] We obtain the residual ε_{icm} after estimating Equation (4). We then estimate the second-stage Equation (Equation (3)), additionally controlling for $\hat{\varepsilon}_{icm}$. Based on Wooldridge (2015), the estimates of β_1 are consistent using this CF method. Because $\hat{\varepsilon}_{icm}$ is predicted in the first stage, the standard errors for the estimates in the second

[59]The use of control function method in a duration model has been justified in Martínez-Camblor et al. (2019).

[60]Note that county fixed effects are implicitly included in the hazard specification through the use of a county-specific baseline hazard.

stage are bootstrapped.

Our first-stage estimate is shown in Table A15 in Appendix 7. We find that there is a very strong first-stage effect, implying a takeup rate of roughly 80% among those made eligible for the program. This estimate is very robust to varying the set of controls that are included in the model.

Table 3 shows our mortality results. The first column shows our baseline hazard without correcting for the endogeneity of NCMS enrollment.[61] We find a significant reduction in the hazard rate of dying of almost 50%. In the second column, we include the control function to control for the potential endogeneity of NCMS enrollment. Our estimate falls to about 30%, suggesting positive selection into the NCMS program. This is consistent with the fact that more educated, and therefore healthier, individuals were more informed about the program in its early years, so they were the first to enroll.[62] The estimate on the first-stage residual is less than 1, which is consistent with the positive selection.

How does this result compare to our aggregate estimates? Given that the one-year mortality rate before NCMS in our sample is 0.2, a 30 % reduction in mortality hazard indicates a 29% reduction in mortality rate after NCMS enrollment.[63] Considering the average takeup rate of 80% of NCMS at the county level, our aggregate estimates suggest a decrease of 22% in aggregate mortality rates for a county with 100 percent NCMS enrollees. This CLHLS result shows a larger mortality reduction among the elderly; this is consistent with the fact that the elderly are more vulnerable to health-service amenable mortality than the younger cohort and therefore benefit more from the health insurance expansion (Nolte and McKee, 2012). At the same time, the very large effect in the CLHLS confirms the smaller but still quite large impact in the aggregate data.

[61]All estimates of the mortality hazard model are reported after taking exponential to represent proportional increases or decreases in mortality rates relative to one.

[62]We provide evidence that early enrollment in NCMS is significantly associated with better health status at the baseline in Appendix 7 Table A16.

[63]Given the estimated hazard ratio of 0.693, we have $\frac{h(t|NCMS\, enrollment_t=1)}{h(t|NCMS\, enrollment_t=0)} = 0.693$. Based on the relationship between mortality hazard h(t) and mortality rate $F(t)$ in $-ln(1-F(t)) = \int_0^t h(u)du$, we have $\frac{-ln(1-F(t|NCMS\, enrollment_t=1)}{-ln(1-F(t|NCMS\, enrollment=0)} = \frac{\int_0^t h(u|NCMS\, enrollment_t=1)du}{\int_0^t h(u|NCMS\, enrollment_t=0)du} = 0.693$. We then back out $F(t|NCMS\, enrollment_t = 1)$ since $F(t|NCMS\, enrollment_t = 0) = 0.2$ in our sample.

The rich data available in the CLHLS allow us to go beyond mortality to look at other health outcomes as well. In particular, we have five other health outcome measures in a sample of survivors. To do so, we move from our hazard framework to linear regression at the individual level:

$$Y_{ict} = \beta_0 + \beta_1 NCMS_{ict} + \mathbf{X_i'}\boldsymbol{\beta} + \delta_{ct} + AgriHukou_i \times \delta_t + AgriHukou_i \times \delta_c + \epsilon_{ict}, \quad (5)$$

where Y_{ict} is a measure of health outcome of individual i in county c in wave t; and $NCMS_{ict}$ indicates that individual i has enrolled in the NCMS in wave t. Because enrolling in the NCMS is voluntary, $NCMS_{ict}$ is endogeneous in Equation (5). We use the IV method to address the endogeneity, and the first stage specification is:

$$NCMS_{ict} = \alpha_0 + \alpha_1 NCMS_{ct} \times AgriHukou_i + \mathbf{X_i'}\boldsymbol{\alpha} + \delta_{ct} + AgriHukou_i \times \delta_t + AgriHukou_i \times \delta_c + \tau_{ict}. \quad (6)$$

Compared Equation (6) with (5), the excluded instrument for $NCMS_{ict}$ is the interaction term between NCMS passage in that county c and individual i's Hukou status ($NCMS_{ct} \times AgriHukou_i$). In both Equations (5) and (6), the vector of covariates, $\mathbf{X_i}$, includes dummies for age, gender, minority status, household size, housing income per capita, drinking, smoking, and an indicator of getting adequate medical service when in serious illness in childhood. These covariates are held constant at their baseline values.

This is a very rich and completely saturated model: we control for county-by-wave fixed effects (δ_{ct}) to account for unobserved time-varying characteristics that are specific to a county, Hukou-by-wave fixed effects ($AgriHukou_i \times \delta_t$) to account for unobserved time-varying characteristics that are specific to a Hukou type, and Hukou-by-county fixed effect ($AgriHukou_i \times \delta_c$) to control for unobserved time-invariant characteristics of individuals with the same Hukou type in a county.[64] By including this full set of interactions in both the first and second stages, we are running a triple-difference IV model.[65] The standard errors are

[64] We are unable to control for $AgriHukou_i \times \delta_t$ and $AgriHukou_i \times \delta_c$ in the hazard anyasis.

[65] Our estimation results remain robust when we additionally control for individual fixed effects, which are reported in Tables A21 - A23 in Appendix 7.

clustered at the county level.

Of course, our health and utilization measures are only measured for those who survive. Since we have shown that the NCMS has an important impact on survival, if those who are marginally kept alive by NCMS are different than the average survivor, it will impart a sample selection bias to our estimates. While we cannot measure the size of this bias, it seems likely that the marginal individuals kept alive by NCMS will be sicker and have higher utilization than the average person still alive. This will bias against finding effects on other health outcomes and towards finding effects on utilization. Fortunately, the very low mortality rate in the CHNS data that we use next allows us to confirm our findings without this bias; the mortality rate in the CHNS is only 0.3% between waves.

The first two outcomes are indicators for being seriously ill over the past two years and for limited activities of daily living. The third variable is mental health, an indicator for a positive state of mind. We also have two dummies of health status: one is self-reported, and the other is interviewer-reported.

The results for these other health measures are shown in the third column of Table 3, using our IV model. We find a significant reduction in the probability of being seriously ill and limitations of activities of daily living, and a significant improvement in mental health; these effects amount to 50% to 70% of the sample mean, in line with the one-third reduction in the mortality hazard we observe in these data. The effects on both self-reported health and reporter-reported health are statistically significant, showing an effect which is roughly 85% as large as the share of the population not in good health. Recall that these effects on non-acute outcomes are not surprising given that that chronic illnesses are frequently treated inpatient in Chinese hospitals, as discussed earlier.

To assess mechanisms, the final column of Table 3 extends our results to examine several measures of healthcare utilization available in the CLHLS, continuing to control for selection. The first is a dummy variable for respondents reporting that they received "adequate medical service". Only 8% of the sample reports that they do not receive adequate medical service, and this gap is closed under NCMS. We then show that the odds of incurring any medical expenditure rise by 11.8 percentage points if individuals are covered by NCMS, which is over 50% of the share of those who report not having positive medical expenditure. Among those with positive expenditure, there is a 42% increase in expenditure.[66]

The CLHLS data, therefore, usefully confirms our finding of significant mortality reductions and shows that these are reflected in other health measures as well. Moreover, the results show a substantial increase in healthcare utilization—consistent with the improvement in health.

6.2 CHNS Results

The other micro-data we use is the CHNS, which is a community-based survey of 225 communities from 54 counties in five waves from 2000 to 2011. We estimate models of the same specification as Equations (5) and (6).[67] Due to data availability, the vector of covariates includes gender, age, minority status, household size, years of schooling, household income per capita, and a set of variables at the community level to measure socioeconomic, transportation, and sanitation conditions.

There are two key differences between the CLHLS and CHNS data to keep in mind in comparing the results. First, the CLHLS data includes only elderly individuals, while the CHNS includes the non-elderly as well. Second, the CLHLS measures medical utilization over the past year, while the CHNS measures it only over the previous four weeks.

Table 4 presents our key results for the CHNS. In the first column, we examine other measures of health outcomes. In particular, we focus on dummy variables for having a cognitive deficit or limitations in activities of daily living, which are measured for those age 55+ only. The odds of a cognitive deficit fall by almost half, while the odds of a limitation

[66]The CLHLS did not ask about the precise medical expenditure until wave 2005. Therefore, we use waves 2005, 2008, 2011 and 2014 to do the analysis.

[67]In particular, we define one's enrollment in NCMS based on the question "which of the following types of medical insurance do you have—cooperative insurance" combined with "Does this village/neighborhood have this type of medical insurance—cooperative insurance", following Lei and Lin (2009). In particular, we first identify village/neighborhood with NCMS adoption based on their reports of having cooperative insurance after 2003. We then define individuals as NCMS enrollees if they report enrolling in cooperative insurance and also live in a village/neighborhood that adopted NCMS. With these two questions, the measure of NCMS enrollment from CHNS is reliable.

in activities of daily living fall by more than half. We also investigate additional health outcomes for the non-elderly population (those below the age of 60). While these estimates are somewhat imprecise, the results in Table A20 of Appendix 7 suggest that enrollment in the NCMS may have reduced the likelihood of experiencing various health issues (such as sickness or injury, fever or cough, headaches, and joint or muscle pain) in the past four weeks.

The second column examines measures of healthcare utilization. We find that healthcare usage rises by 8.4 percentage points, corresponding to a 77% increase compared to the sample average propensity of 0.109. This is composed of a 1.9 percentage points rise in the use of inpatient care and a 7.2 percentage points rise in outpatient care. The survey also asks separately about the receipt of preventive care, and this rises by 5.4 percentage points. Total medical expenditure increases by 39% among those with positive expenditure.[68] But the estimate is not precise due to the relatively small sample size.

Therefore, both of these valuable micro-data surveys confirm the transformational results of the NCMS program in a treatment-on-the-treated framework.[69] There are dramatic improvements in healthcare utilization, as well as improvements in health outcomes along a wide variety of dimensions. We also investigate the influence of NCMS enrollment on individuals' overall well-being, particularly their self-reported life satisfaction. Although the estimation is not precise, the findings presented in Table A19 of Appendix 7 might encompass various benefits associated with health insurance, such as decreased stress levels or reduced social stigma.

We acknowledge the potential for spillover effects, such as changes in insurance coverage or expenditures within a family, or supply-side changes induced by policy. Spillover within families could enable both members with agricultural hukou, who are not enrolled in the program, and members with urban hukou, who are ineligible for the program, to benefit

[68]There might be concerns regarding potential sample selection given that the sample is restricted to individuals with positive medical spending. However, given this condition, we can explore the effect of NCMS enrollment on an individual's medical expenditure at the intensive margin. This complements our analysis of healthcare usage at the extensive margin.

[69]Similar to our assessment of the identification assumption for Equation (1) using the aggregate data in Appendix 5, we also examine the research design using micro-level data. We demonstrate that the interaction between NCMS adoption and the baseline agricultural share is uncorrelated with the previous county variables or their involvement in recent years, as shown in Table A17 and A18 of Appendix 7.

from the adoption of the NCMS. As our IV estimation assumes that the health effects come from insurance coverage only, it may overstate impacts if there are spillovers. However, given the first-stage estimates presented in Table A15 in Appendix 7, the reduced-form estimates are approximately 70%-80% the size of the IV estimates. These spillover effects do not significantly threaten our positive findings concerning the health effects. In the following section, we explore the role of policy-induced supply responses in the health effects of the NCMS.

7 Potential Mechanisms

7.1 Out-of-Pocket Expenditure

A main mechanism through which our results could be operating is reduced OOP expenditure exposure for NCMS enrollees. A decrease in OOP expenditure due to NCMS enrollment may alleviate an individual's concerns about future health expenditure risk, potentially improving their mental health. Furthermore, reduced OOP expenditure might free up resources for other forms of consumption, such as food, which could in turn contribute to better health outcomes. For instance, Bai and Wu (2014) find that NCMS coverage led to an increase in non-medical-related consumption by more than 5% among rural households.

Table 5 shows the impact of the NCMS on OOP expenditure, using our treatment-on-the-treated framework in both the CLHLS and CHNS data. The first column uses a dependent variable from a question asked in the CLHLS about whether the respondent refuses any necessary medical service due to financial difficulty. We find a highly significant reduction in this measure of 5.1 percentage points, which is 85% of the population mean. Unfortunately, the CLHLS does not ask about OOP expenditure, but the CHNS does.

The second column shows that, in fact, we find an insignificant effect on (the log of) OOP expenditure. This is surprising given the insurance protection provided by the program, but it also reflects the imprecision in our estimate. This is because only 330 persons in our data have OOP expenditure in the month before the survey, both before and after the NCMS becomes available.[70]

In addition, the drop in OOP expenditure is smaller than what might be expected since total medical expenditure is rising so rapidly. The third column changes the dependent variable to the ratio of OOP expenditure to total expenditure. This shows a significant reduction of 29 percentage points in the share of medical expenditure borne out of pocket.

Finally, the NCMS program is clearly reducing the risk of extreme levels of OOP expenditure. Figure 5 shows the relationship between total medical expenditure and OOP expenditure before and after enrollment in NCMS. Before the NCMS, there is an essentially 1-1 correspondence between total expenditure and OOP expenditure.[71] But after NCMS, there estimated a positive relationship but much below 1, so that for those with the highest total expenditure, they are bearing a much smaller share of OOP expenditure.

7.2 Supply Expansions

Before 2000, the Chinese healthcare market was almost exclusively run by the public sector. Although private clinics were allowed to operate, they were extremely limited through unfavorable tax and reimbursement policies, and private investment in public medical facilities was illegal. In 2003, the government began allowing private investment in public health facilities in order to address the coming increase in the demand for medical care. Private capital was encouraged to invest in the public healthcare sector via forms such as entrusted operation and joint venture. Private investors were allowed to send delegates to the administrative boards of public facilities (although board leadership was still public); these delegates would take part in operations of public facilities and may propose to hire professional managers to improve the management efficiency.

This has two potential implications for our study. The first is that our results could be driven by contemporaneous changes in supply that happen to be correlated with NCMS expansions. The second is that some of our results could be explained by induced supply-side

[70]Though implementing such a restriction on the sample may introduce sample selection issues, it is necessary to facilitate the comparison of out-of-pocket spending distributions before and after the adoption of the NCMS.

[71]This fact confirms that the charity care or uncompensated care was rare in China.

effects, as in Gruber et al. (2014). By raising demand for care, public programs can expand the supply of care available. The Chinese context offers a potential multiplier effect for this channel through the simultaneous liberalization of private investment in public healthcare facilities.

To address these concerns, we include controls for supply in our outcome models. The results are shown in Table 6, controlling for two measures of supply: the number of healthcare staff and the number of hospital beds per 1,000 people. In fact, the coefficients on supply are highly significant, and indicate that increases in supply did lead to reductions in mortality and increases in life expectancy. In particular, we find that a 10% rise in healthcare supply lowered mortality by around 3%, and raised life expectancy by around 0.6%.

But we find that including these supply measures has little impact on our policy variables, which fall by only 8-10%. This helps address concerns that simultaneous supply expansions are driving our results. It also puts an upper bound on the amount of our policy effects that are due to induced supply responses.[72]

8 Conclusions

The rise of modern medicine in the developing world has brought with it the need to address high medical costs relative to individual resources. A natural approach to addressing these costs is through expanded public insurance systems. A variety of countries around the world have introduced such systems in recent decades, with the most notable recent addition being the PM-JAY program in India which is ultimately designed to cover the bottom 40% of the population, or almost 500 million persons.

The largest of these expansions in terms of persons covered is the NCMS program introduced in rural China in 2003. Previous research on this program has delivered mixed and generally muted evidence of success from this program, particularly in terms of health outcomes. Our research using aggregate health data delivers a very different conclusion: the NCMS was an unqualified success. The program saved more than one million lives per

[72]Mathematically, this is equivalent to regressing supply on our policy variable and multiplying by the change in policy to estimate the impact of policy-induced supply on mortality.

year, explaining nearly four-fifths of the rise in life expectancy during its first eight years. Moreover, the program was highly cost effective, with a cost per life saved well below common benchmarks.

Given these striking results from aggregate data, we explore the leading sources of health micro-data for China to confirm and extend our findings. We find from two different surveys that the NCMS program induced major increases in healthcare utilization, as well as improvements in a variety of other health measures, ranging from self-reported health to activities of daily living.

Finally, we explore the mechanisms through which NCMS resulted these effects. We show modest reductions in OOP expenditure overall, but very sizeable reductions at the top of the medical expenditure distribution. We suggest that at most a modest share of the effect arises through induced supply responses, and that there is little evidence of ex-ante moral hazard effects of the new coverage.

The fact that this incredible health policy success has gone largely underappreciated by the health policy community is striking and suggests the potential value of revisiting the impacts of major health expansions in the developing world. Indeed, these results are quite consistent with findings for the health benefits of the Affordable Care Act (ACA) in the US. Miller et al. (2021) find that, among individuals aged 19-64, expanded Medicaid saved one life per 595 people insured; among those 55-64, it was one life per 97 people insured. Goldin et al. (2021) find that among those aged 45-64, there is one life saved per 52 people enrolled. Our results imply that there is one life saved per 900 persons enrolled, which is actually somewhat weaker than the results from these US-based studies. Comparable well-identified studies across a broader set of nations could allow us to understand under which conditions health insurance expansions can be most effective.

That said, there remain many questions to answer about the NCMS. In particular, there is much more to do to understand the heterogeneity of these effects, and how it depended on local decisions of features such as the nature of outpatient coverage. Moreover, as China urbanizes, there are questions about the ongoing value of programs that are differentiated between urban and rural areas; comparable analysis of the urban insurance expansion that began in 2008 would help inform this issue.

References

Agha, L. and D. Zeltzer (2022). Drug diffusion through peer networks: The influence of industry payments. *American Economic Journal: Economic Policy 14*(2), 1–33.

Bai, C.-E. and B. Wu (2014). Health insurance and consumption: Evidence from China's New Cooperative Medical Scheme. *Journal of Comparative Economics 42*(2), 450–469.

Blumenthal, D. and W. Hsiao (2015). Lessons from the East-China's rapidly evolving health care system. *New England Journal of Medicine 372*(14), 1281–1285.

Burns, L. R. and G. G. Liu (2017). *China's healthcare system and reform.* Cambridge University Press.

Chen, Y. and G. Z. Jin (2012). Does health insurance coverage lead to better health and educational outcomes? Evidence from rural China. *Journal of Health Economics 31*(1), 1–14.

Cheng, L., H. Liu, Y. Zhang, K. Shen, and Y. Zeng (2015). The impact of health insurance on health outcomes and spending of the elderly: Evidence from China's New Cooperative Medical Scheme. *Health Economics 24*(6), 672–691.

Curtin, L. R. and R. J. Klein (1995). Direct standardization (age-adjusted death rates): US Department of Health and Human Services. *Public Health Service, Centers for Disease Control and Prevention, National Center for Health Statistics 6*.

De Chaisemartin, C. and X. D'Haultfoeuille (2022a, March). Difference-in-differences estimators of intertemporal treatment effects. Working Paper 29873, National Bureau of Economic Research.

De Chaisemartin, C. and X. D'Haultfoeuille (2022b, 06). Two-way fixed effects and differences-in-differences with heterogeneous treatment effects: A survey. *The Econometrics Journal*. utac017.

Erten, B. and J. Leight (2021). Exporting out of agriculture: The impact of WTO accession on structural transformation in China. *Review of Economics and Statistics 103*(2), 364–380.

Feng, J. (2013). China's health transition and future health care financing. In *Unfinished Reforms in the Chinese Economy*, pp. 229–267. World Scientific.

Gertler, P., L. Locay, and W. Sanderson (1987). Are user fees regressive? The welfare implications of health care financing proposals in Peru. *Journal of Econometrics 36*(1-2), 67–88.

Goldin, J., I. Z. Lurie, and J. McCubbin (2021). Health insurance and mortality: Experimental evidence from taxpayer outreach. *Quarterly Journal of Economics 136*(1), 1–49.

Goodman-Bacon, A. (2021). The long-run effects of childhood insurance coverage: Medicaid implementation, adult health, and labor market outcomes. *American Economic Review 111*(8), 2550–2593.

Gruber, J. (1994). The incidence of mandated maternity benefits. *The American Economic Review*, 622–641.

Gruber, J., N. Hendren, and R. M. Townsend (2014). The great equalizer: Health care access and infant mortality in Thailand. *American Economic Journal: Applied Economics 6*(1), 91–107.

Hammitt, J. K. and Y. Zhou (2006). The economic value of air-pollution-related health risks in China: A contingent valuation study. *Environmental and Resource Economics 33*(3), 399–423.

Huang, W. and H. Liu (2023). Early childhood exposure to health insurance and adolescent outcomes: Evidence from rural china. *Journal of Development Economics 160*, 102925.

Huang, W. and C. Zhang (2021). The power of social pensions: Evidence from China's New Rural Pension Scheme. *American Economic Journal: Applied Economics 13*(2), 179–205.

Lei, X. and W. Lin (2009). The New Cooperative Medical Scheme in rural China: Does more coverage mean more service and better health? *Health Economics 18*(S2), S25–S46.

Litvack, J. I. and C. Bodart (1993). User fees plus quality equals improved access to health care: Results of a field experiment in Cameroon. *Social Science & Medicine 37*(3), 369–383.

Liu, K. (2016). Insuring against health shocks: Health insurance and household choices. *Journal of Health Economics 46*, 16–32.

Long, Q., R. Klemetti, Y. Wang, F. Tao, H. Yan, and E. Hemminki (2012). High Caesarean section rate in rural China: Is it related to health insurance (New Co-operative Medical Scheme)? *Social Science & Medicine 75*(4), 733–737.

Martínez-Camblor, P., T. A. MacKenzie, D. O. Staiger, P. P. Goodney, and A. James OMalley (2019). An instrumental variable procedure for estimating Cox models with non-proportional hazards in the presence of unmeasured confounding. *Journal of the Royal Statistical Society: Series C (Applied Statistics) 68*(4), 985–1005.

McDermott, K. W. and M. Roemer (2021). Most frequent principal diagnoses for inpatient stays in US hospitals, 2018: Statistical brief# 277.

Milcent, C. (2018). *Healthcare reform in China: From violence to digital healthcare*. Springer.

Miller, S., N. Johnson, and L. R. Wherry (2021). Medicaid and mortality: New evidence from linked survey and administrative data. *Quarterly Journal of Economics 136*(3), 1783–1829.

Neumark, D., J. I. Salas, and W. Wascher (2014). Revisiting the minimum wage¡ªemployment debate: Throwing out the baby with the bathwater? *Ilr Review 67*(3_suppl), 608–648.

Nolte, E. and C. M. McKee (2012). In amenable mortality—deaths avoidable through health care—progress in the US lags that of three European countries. *Health Affairs 31*(9), 2114–2122.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics 37*(2), 187–204.

Souteyrand, Y. P., V. Collard, J. P. Moatti, I. Grubb, and T. Guerma (2008). Free care at the point of service delivery: A key component for reaching universal access to HIV/AIDS treatment in developing countries. *AIDS 22*, S161–S168.

Sun, L. and S. Abraham (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics 225*(2), 175–199.

Sun, X., S. Jackson, G. Carmichael, and A. C. Sleigh (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.

Wagstaff, A., M. Lindelow, G. Jun, X. Ling, and Q. Juncheng (2009). 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.

Wang, H. and J. He (2014). Estimating the economic value of statistical life in China: A study of the willingness to pay for cancer prevention. *Frontiers of Economics in China 9*(2), 183–215.

Wolfers, J. (2006). Did unilateral divorce laws raise divorce rates? A reconciliation and new results. *American Economic Review 96*(5), 1802–1820.

Wooldridge, J. M. (2015). Control function methods in applied econometrics. *Journal of Human Resources 50*(2), 420–445.

Yip, W. (2020). *Health Care System Reform and Policy Research in China*. World Scientific.

Yip, W. and W. Hsiao (2008). The Chinese health system at a crossroads. *Health affairs 27*(2), 460–468.

Yip, W. and W. Hsiao (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.

Zhou, M., S. Liu, M. Kate Bundorf, K. Eggleston, and S. Zhou (2017). Mortality in rural China declined as health insurance coverage increased, but no evidence the two are linked. *Health Affairs 36*(9), 1672–1678.

Zombré, D., M. De Allegri, and V. Ridde (2017). Immediate and sustained effects of user fee exemption on healthcare utilization among children under five in Burkina Faso: A controlled interrupted time-series analysis. *Social Science & Medicine 179*, 27–35.

Figures

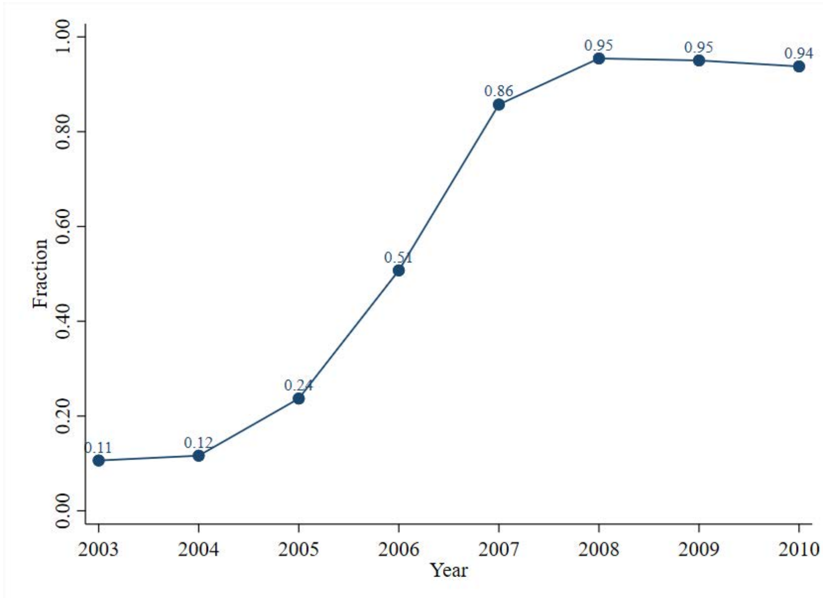


Figure 1: NCMS Rollout

Notes: The figure plots the fraction of counties implementing the NCMS over time. The number of NCMS counties is obtained from China Health Statistical Yearbook (2004-2011) and the total number of counties from 2011 China Civil Affairs' Statistical Yearbook.

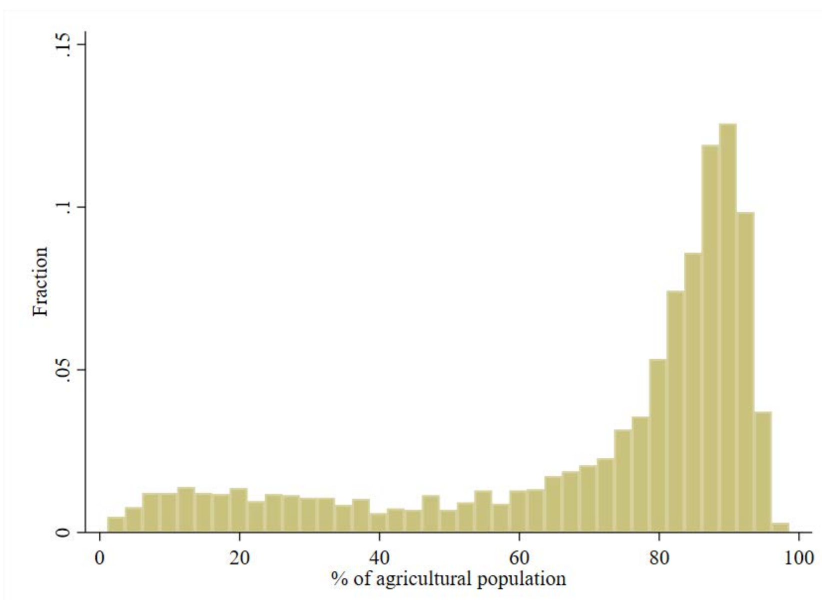
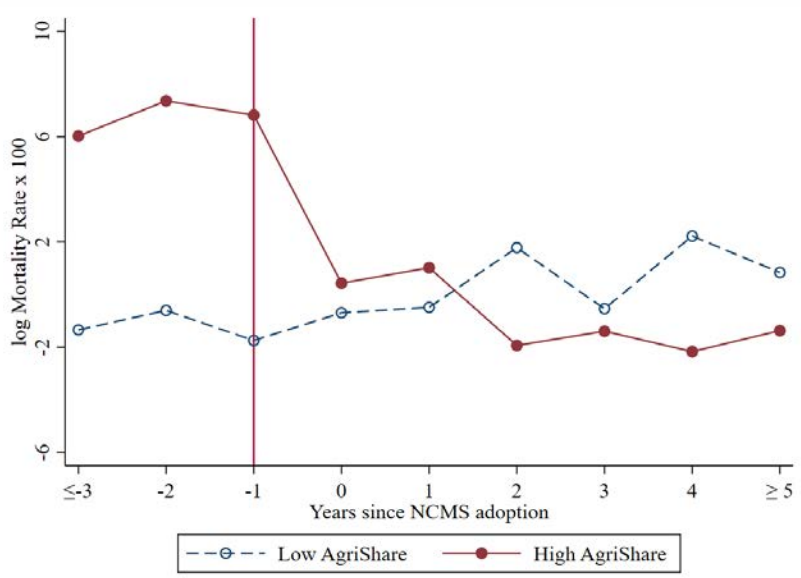


Figure 2: Distribution of the share of agricultural Hukou population across counties

Notes: The figure plots the distribution of the share of the population with an agricultural Hukou across counties using China's 2000 Population Census.

Panel A. Age-adjusted mortality rate



Panel B. Life expectancy at birth

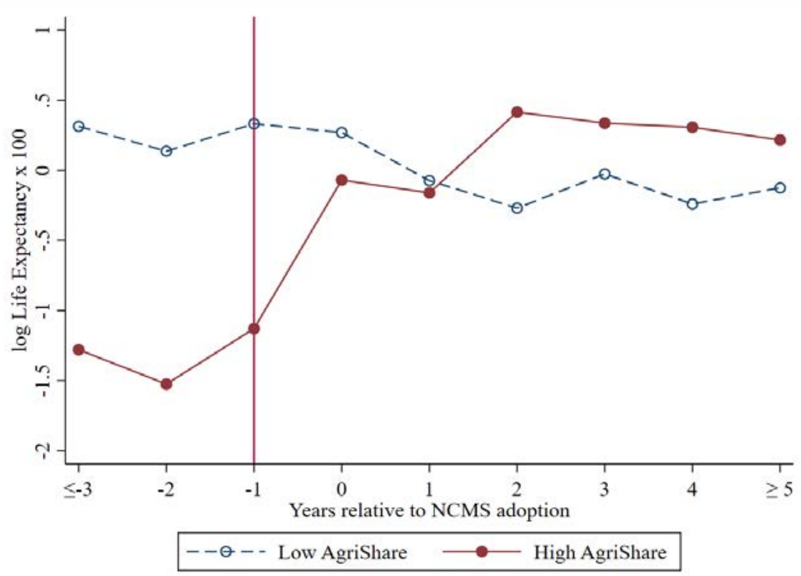
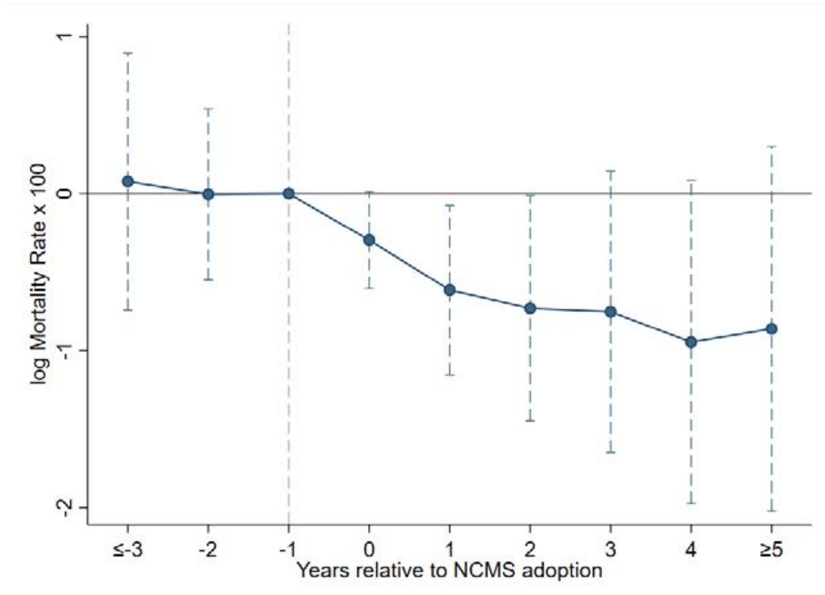


Figure 3: Mortality rate and life expectancy around NCMS adoption: subsamples by agricultural share

Notes: The figures plot the average value of the residuals from regressing mortality rate (Panel A) and life expectancy (Panel B) on NCMS-timing-by-year FE and county-specific linear trend. The regressions are weighted by county population. Counties in the sample are split by the median value of the share of the population with an agricultural Hukou. Observations more than three years (five years) before (after) the NCMS are binned into groups. The sample period is from 2004 to 2010.

Panel A. Age-adjusted mortality rate



Panel B. Life expectancy at birth

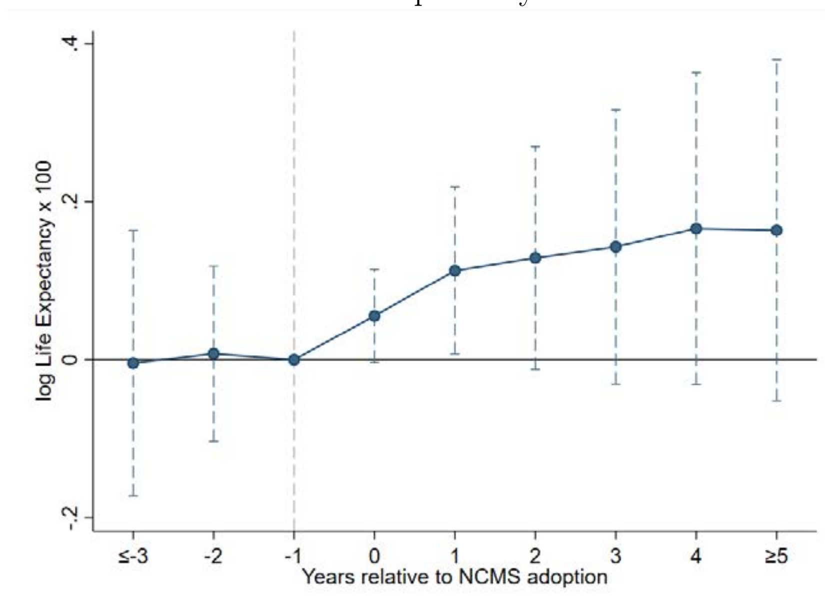


Figure 4: Mortality rate and life expectancy around NCMS adoption

Notes: The figure plots the estimated coefficients on the interactions between year-to-NCMS dummies and 2000 agricultural share from the regression model specified in Equation (2). The year before NCMS adoption is omitted, so the estimates are normalized to zero in that year. Observations more than 3 years (5 years) before (after) the NCMS are binned into groups. Results are weighted by county population. Sample period is from 2004 to 2010. The dashed lines represent the 95 percent confidence intervals based on standard errors clustered at the county level.

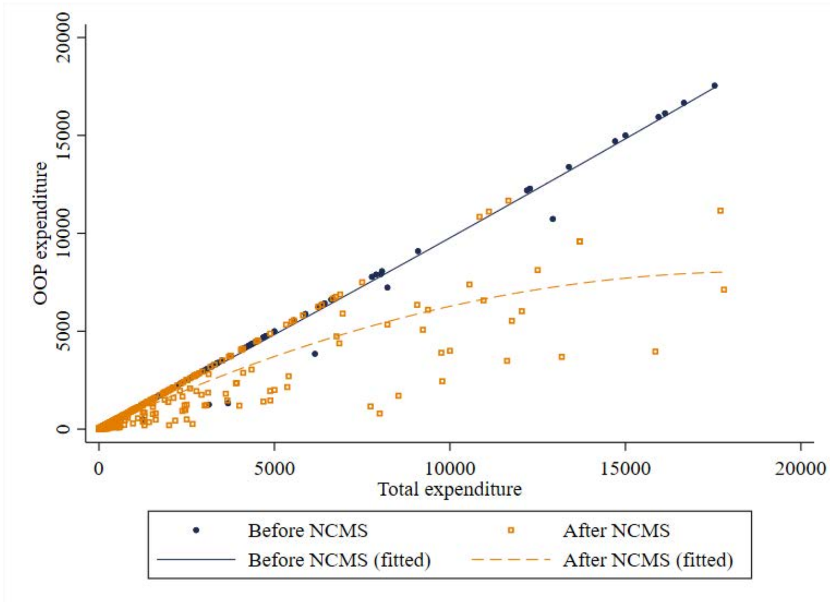


Figure 5: OOP expenditure vs. total expenditure by NCMS enrollment

Note: The figure plots the OOP expenditure against the total expenditure from the CHNS. The sample is restricted to those with positive medical expenditure both before and after enrolling in the NCMS.

Tables

Table 1: Summary of public health insurance programs in developing countries

Year	Country	Name of Insurance Scheme	Number of People (year)
1990	Brazil	Sistema Único de Saúde (SUS)	200 million (2018)
1992	Vietnam	Social Health Insurance	80 million (2018)
1993	Colombia	Regimen Subsidiado (SR)	34 million (2016)
1995	Philippines	Philippines Health Insurance Corporation (PhilHealth)	up to 104 million (2018)
1999	Rwanda	Mutuelles de Santé	12 million (2018)
2002	Thailand	Universal Coverage Health Scheme (UCS)	53 million (2018)
2003	China	New (Rural) Cooperative Medical Scheme (NCMS or NRCMS)	836 million (2010)
2005	Ghana	National Health Insurance Scheme (NHIS)	11 million (2014)
2018	India	Pradhan Mantri Jan Arogaya Yojana (PM-JAY)	103 million (2019)

Notes: The table shows information about new health insurance programs in some of the largest developing countries over the past 30 years, based on authors' review of material on health insurance expansions in developing country. Data sources are listed in Appendix 1. For the Philippines, there is larger controversy over the reach of the program, so we list an upper bound here.

Table 2: Impacts of NCMS enrollment on mortality rates and life expectancy using DSP data

	(1) Baseline	(2) Controlling for local government inputs	(3) Removing economic covariates	(4) Sample period 2004 to 2007	(5) Dropping outliers	(6) Placebo tests
Panel A: Age-adjusted mortality rate						
Post NCMS x AgriShare2000	-0.197**	-0.186**	-0.197**	-0.220**	-0.210**	-0.038
	(0.091)	(0.091)	(0.091)	(0.104)	(0.093)	(0.112)
Adjusted R-squared	0.803	0.803	0.803	0.768	0.803	0.786
Mean	546.30	546.30	546.30	550.91	575.1	822.28
Panel B: Life expectancy at birth						
Post NCMS x AgriShare2000	0.040**	0.038**	0.039**	0.046**	0.042**	-0.011
	(0.018)	(0.018)	(0.019)	(0.021)	(0.019)	(0.026)
Adjusted R-squared	0.740	0.740	0.740	0.656	0.740	0.768
Mean	75.35	75.35	75.35	75.28	74.51	71.00
County FE	Yes	Yes	Yes	Yes	Yes	Yes
Province-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
Agricultural share x year FE	Yes	Yes	Yes	Yes	Yes	Yes
NCMS-timing-by-year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-specific linear trend	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes
Obs	847	847	847	484	819	678

Notes: The table shows the estimated coefficients from Equation (1) in various specifications. Panel A shows the results on age-adjusted mortality rates and Panel B shows those on life expectancy at birth. Column (1) reports the baseline results from Equation (1) using DSP sample from 2004 to 2010. Columns (2) includes variables of other local policies, such as the number of social aid centers and investment in fixed asset per capita. Column (3) reports results from specifications removing economic covariates. Columns (4) restricts the sample period to 2004 -2007. Column (5) drops the four counties in the top and bottom 1% of our residual distribution. Columns (6) reports the results from a placebo test using the sample from 1994 to 2000. Age-adjusted mortality rate and life expectancy are in logarithm. Results are weighted by county population. Standard errors clustered at the county level are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 3: Impacts of NCMS enrollment on health outcomes and healthcare utilization using CLHLS

	Mortality hazard		Other health outcomes	Healthcare Utilization
	(1)	(2)	(3)	(4)
	Baseline	Control function		
NCMS enrollment	0.553***	0.693***		
	(0.025)	(0.065)		
1^{st}-stage Residuals		0.776***		
		(0.073)		
Obs	78,446	78,446		
Being seriously ill			-0.129**	
			(0.063)	
Obs			53,224	
Mean			0.183	
Limited ADL			-0.125**	
			(0.063)	
Obs			53,224	
Mean			0.252	
Mental health			0.111**	
			(0.054)	
Obs			53,224	
Mean			0.228	
Self-reported health status			0.123**	
			(0.051)	
Obs			53,224	
Mean			0.853	
Interviewer-reported health status			0.122**	
			(0.050)	
Obs			53,224	
Mean			0.862	

Continued on next page

Get adequate medical service				0.089**
				(0.043)
Obs				58,565
Mean				0.920
Positive medical expenditure				0.118*
				(0.067)
Obs				31,378
Mean				0.791
Total medical cost (log)				0.418*
				(0.222)
Obs				24,832
Mean				2.733
Year FE	Yes	Yes	No	No
County stratification	Yes	Yes	No	No
Baseline covariates	Yes	Yes	Yes	Yes
County-by-wave FE	No	No	Yes	Yes
Hukou-by-wave FE	No	No	Yes	Yes
Hukou-by-county FE	No	No	Yes	Yes

Notes: The table reports the estimation results from Equations (3) - (6) using the CLHLS. Column (1) shows how NCMS enrollment is related to one's mortality hazard based on Equation (3). Column (2) presents results from the control function method, with county NCMS $\times$ AgriHukou as an instrument for NCMS enrollment. All estimates are reported after taking exponential to represent proportional increases or decreases in mortality rates relative to one. Standard errors in parentheses are computed based on 1,000 bootstraps and converted via the delta method. Columns (3) and (4) report the results on other health outcomes and healthcare utilization in the sample of survivors. Each cell presents the IV coefficients on NCMS enrollment based on Equations (5) and (6). Dependent variables in these regressions are shown in the first column. Information on medical expenditure is available after 2005. Total medical expenditure is in 1,000. Standard errors clustered at the county level are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 4: Impacts of NCMS enrollment on health outcomes and healthcare utilization using CHNS

	(1) Health outcomes	(2) Healthcare utilization	(3) Medical expenditure
Cognitive deficit	-0.220*		
	(0.126)		
Obs	5,942		
Mean	0.509		
Limited ADL	-0.257**		
	(0.128)		
Obs	5,942		
Mean	0.333		
Seek medical care		0.084**	
		(0.040)	
Obs		46,519	
Mean		0.109	
Inpatient care		0.019*	
		(0.010)	
Obs		46,519	
Mean		0.009	
Outpatient care		0.072**	
		(0.035)	
Obs		46,519	
Mean		0.095	
Preventive care		0.054**	
		(0.026)	
Obs		46,519	
Mean		0.037	

Continued on next page

Total medical expenditure			0.387
(log)			
			(0.638)
Obs			4,005
Mean			1.537
Covariates	Yes	Yes	Yes
County-by-wave FE	Yes	Yes	Yes
Hukou-by-wave FE	Yes	Yes	Yes
Hukou-by-county FE	Yes	Yes	Yes

Notes: The table shows the results from Equations (5) and (6) about health outcomes, healthcare utilization, and medical expenditure using CHNS. Each cell presents the IV coefficients on NCMS enrollment based on Equations (5) and (6). Dependent variables in these regressions are shown in the first column. Limited ADL and cognitive deficit are measured for those age 55+ only, and the information is available in waves 2000 - 2006. Sample in column (3) is restricted to those with positive medical expenditure. Total medical expenditures are in 1,000. Standard errors clustered at the county level are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 5: Potential mechanisms: financial protection

	(1) Financial difficulty	(4) ln(OOP expenditure)	(5) OOP/total expenditure
NCMS enrollment	-0.051**	-0.429	-0.289**
	(0.026)	(1.163)	(0.135)
Covariates	Yes	Yes	Yes
County-by-wave FE	Yes	Yes	Yes
Hukou-by-wave FE	Yes	Yes	Yes
Hukou-by-county FE	Yes	Yes	Yes
Obs	52,046	680	680
Mean	0.059	1.112	0.952

Notes: The table reports the estimation results from Equations (5) and (6) and the dependent variables are shown in the first row. Column (1) uses the CLHLS sample and financial difficulty is an indicator for not seeking any medical service due to lack of money. This information is available after 2000. Columns (2) and (3) use the CHNS sample and the sample is restricted to those with positive OOP expenditure both before and after the NCMS initiation in their counties. Expenditures are in 1,000. Standard errors clustered at the county level are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table 6: Potential mechanisms: healthcare supply expansions

	(1)	(2)
Panel A. Age-adjusted mortality rates		
Post NCMS x AgriShare	-0.180**	-0.178**
	(0.089)	(0.089)
Healthcare staff (log)	-0.283***	
	(0.099)	
Hospital beds (log)		-0.354***
		(0.131)
Mean	546.23	546.23
Panel B. Life expectancy at birth		
Post NCMS x AgriShare	0.036**	0.036**
	(0.018)	(0.018)
Healthcare staff (log)	0.058**	
	(0.024)	
Hospital beds (log)		0.064**
		(0.025)
Mean	75.35	75.35
County FE	Yes	Yes
Province-by-year FE	Yes	Yes
AgriShare2000 x year FE	Yes	Yes
NCMS-timing-by-year FE	Yes	Yes
County-specific linear trend	Yes	Yes
Covariates	Yes	Yes
Obs	847	847

Notes: The table shows the results of how NCMS adoption and local healthcare supply jointly affect health outcomes age-adjusted mortality rate (Panel A) and life expectancy at birth (Panel B). Columns (1) and (2) show the OLS results of Equation (1), additionally controlling for local healthcare staff and hospital beds (log), respectively. Healthcare supply is measured as the logarithm of the number per 1,000 people. Results are weighted by county population. Standard errors clustered at the county level are in parentheses. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.