I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13932

Pension Policies, Retirement and Human Capital Depreciation in Late Adulthood

Plamen Nikolov
Alan Adelman

DECEMBER 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13932

Pension Policies, Retirement and Human Capital Depreciation in Late Adulthood

Plamen Nikolov
State University of New York, IZA and Harvard Institute for Quantitative Social Science

Alan Adelman
State University of New York

DECEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Pension Policies, Retirement and Human Capital Depreciation in Late Adulthood*

Economists have mainly focused on human capital accumulation and considerably less on the causes and consequences of human capital depreciation in late adulthood. Studying human capital depreciation over the life cycle has powerful economic consequences for decision-making in old age. Using data from China, we examine how a new retirement program affects cognitive performance. We find large negative effects of pension benefits on cognitive functioning among the elderly. We detect the most substantial impact of the program on delayed recall, a significant predictor of the onset of dementia. We show suggestive evidence that the program leads to larger negative impacts among women. We demonstrate that retirement and access to a retirement pension plan plays a significant role in explaining cognitive decline at older ages.

JEL Classification: O12, J24, J26, H55, H75, O15

Keywords: life cycle, cognitive functioning, cognition, aging, health, mental retirement, middle-income countries, LMICs, developing countries, China

Corresponding author:
Plamen Nikolov
Department of Economics
State University of New York (Binghamton)
4400 Vestal Parkway East
Binghamton, NY 13902
USA

E-mail: pnikolov@post.harvard.edu

* We thank Steve Yeh, Charlotte Williams, Matthew Bonci, William Lombardo, Ariunzaya Oktyabri, Sharon Itin, Hongyi Duan, Declan Levine, Xu Wang, and Jian Deng for outstanding research support with this project. We thank David Cutler, Jonathan Kolstad, Eric Edmonds, Daniel Millimet, Carly Urban, Paul Novosad, Zoe McLaren, Carly Trachtman, Brett Matsumoto, Darren Kauffman, Dan Lee, Wei Huang, Shelly Lundberg, Livia Montana, Susan Wolcott, Evan Riehl, Nusrat Jimi, Subal Kumbhakar, Solomon Polachek, Petra Moser, James MacKinnon, and Morten Nielsen for constructive feedback and helpful comments.

I. Introduction

Cognitive abilities are one dimension of human capital, along with education, health, and noncognitive skills. Historically, economists have mainly focused on human capital accumulation (Heckman 2000) and considerably less on the causes and consequences of human capital depreciation. However, recent neuropsychological evidence suggests that the human brain is malleable and open to enhancement even in late adulthood. Cognitive aging is an important and complex phenomenon, and its economic or policy causes are not well understood. This paper examines the effects of a pension program on early retirement and cognitive performance in old age.

China introduced the New Rural Pension Scheme (NRPS) in 2009 to ease demographic pressures and concerns about old-age poverty facing the country (Dorfman et al. 2013).[1,2] In this paper, we examine the causal effect of individual participation in the NRPS program on human capital depreciation among individuals aged 60 and over. Pensioners who reach age 60 and contribute to the NRPS start receiving their pension, comprising a basic pension from the government and a portion determined from their prior contributions. We examine the pension program's effect on cognitive performance in old age. We focus on the link between early retirement, pension benefits, and cognitive decline in old age.

Studying how human capital depreciates over the life cycle has powerful economic consequences. At the micro-level, cognitive functioning is crucial for decision-making as it influences the ability to process information. Elderly individuals are increasingly required to

[1] Feldstein and Liebman (2002) and Cutler and Johnson (2004) overview social insurance programs in developed countries.
[2] The primary factors that precipitated the introduction of the program were demographic and economic challenges: population aging (Bloom and McKinnon 2014), a large rural fraction of the population, rising income inequality (Sabates-Wheeler and Koettl 2010), and weak local institutions to support social protection on their own. The program was financed from commingled funds: local and federal funds financed the first source of the program cost; the second source came from individual contributions. The central government subsidizes 100 percent of the program cost in provinces with low fiscal capacity. In contrast, federal subsidies constitute only 50 percent of the total funding in wealthier provinces with high fiscal capacity.

make complex financial, health, and long-term-care decisions, with significant consequences.[3] Understanding the causes of cognitive decline is also crucial for policy—the relationship between cognitive aging and productivity impacts long-term economic growth. Examining the effect on cognition for the older population in a country such as China may be especially important given the country's lack of intermediary market institutions to assist with financial decisions related to income security or health care provision.

Our empirical estimation relies on the staggered implementation of the NRPS program across rural parts of China between 2009 and 2013. We use a difference-in-difference-in-differences methodology (or triple-difference/DDD estimator)[4] to identify the causal effect of access to the retirement NRPS plan on cognition among aging adults. We rely on identifying variation from three distinct sources: the program rollout at the municipality[5] level (known as *shequs*), the timing of program adoption by each municipality, and that only program participants aged 60 and over received the NRPS retirement income. Our analysis relies on a new data source—the Chinese Health and Retirement Longitudinal Survey (CHARLS)— nationally representative of individuals ages 45 and above. The CHARLS, a sister survey of the U.S.-based Health and Retirement Survey (HRS), directly tests cognition, focusing on two critical cognitive domains: *episodic memory* and components of *intact mental status*. Episodic memory captures fluid intelligence aspects, whereas the mental intactness measure captures both fluid and crystallized intelligence.

Our paper finds clear evidence of adverse effects on cognitive performance in old age among NRPS participants. First, the NRPS program has a significantly negative effect on

[3] Cognitive ability remains an important factor associated with healthy aging. Cognitive decline among aging adults can negatively influence investment behavior and have negative implications for the financial wellbeing in retirement (Korniotis and Kumar 2011).
[4] Similar to Gruber (1994).
[5] Because our primary data source refers to these administrative units as sheques (社区) or communities. Therefore, we refer to these units as *communities* (or *municipalities*) from now on.

cognition among individuals aged 60 or above. Retirement programs are introduced and geared to ensure adults' welfare in old age (Cutler and Johnson 2004). Nevertheless, we provide strong evidence for a clear case on how a new pension program can lead to some unintended and significant adverse consequences for program participants. Specifically, we find that the provision of pension benefits negatively influences immediate recall, delayed recall, and total word recall. Lower performance on the delayed recall memory measure has been a highly accurate detector of dementia among elderly individuals (Welsh et al. 1991).[6] For the total word recall outcome, relative to the cognitive performance on tests before access to program benefits, the estimated effect size for the cognitive decline is 12 percent of a standard deviation (or approximately five percent of the average baseline score on the cognition measure). This result is associated with approximately four years of program exposure. We can benchmark our effect size estimates to general ability measures: a 5-percent drop in the average total recall score due to access to an NRPS pension plan is approximately equivalent to a decline in general intelligence by 1.7 percent (relative to the general population).[7] Our main results are robust to several different specifications that test the validity of the proxy measure of NRPS participation. Furthermore, we examine if longer program exposure exacerbates the cognitive decline, and we find corroborating evidence to substantiate this link.

Earlier retirement can influence cognition among aging adults via several channels. The combination of having both guaranteed retirement income upon reaching age 60 and generous government subsidies for particular contribution levels could boost some permanent income. Consequently, this income boost could reduce incentives for full labor market participation on

[6] It is worth noting that although we find a decline in the delayed recall memory due to participation in the pension program, this does not necessarily imply a greater incidence of dementia due to pension program participation.
[7] Ackerman et al. (2005) investigate the relationship between recall memory measures and general intelligence (fluid intelligence). Ackerman et al. (2005) find that a 1-percent decline in word scores leads to a 0.33-percent decrease in proxies of general intelligence based on a meta-analysis study.

various intensive margins (i.e., reduced effort, reduced hours, and reduced work schedules). Furthermore, and only among those who fully retire, labor force withdrawal due to participation in the NRPS could generate additional benefits: reduced stress, improved personal diets, and improved overall sleep patterns. However, the program could create unintended adverse effects. For example, reducing labor activities could reduce engagements in social activities and worsen mental acuity fitness. The net effect of the NPRS is, therefore, theoretically ambiguous. Although we find that NRPS improves various health behaviors, sleep patterns, and nutritional diets, our analysis of potential mechanisms shows that the program also leads to a substantial reduction in social engagement, volunteering, and activities involving the use of mental capacity. Therefore, given the net adverse effects on the cognition outcomes, the negative effect of the NRPS on social engagements and participation in activities related to mental fitness likely outweighs the positive impact on nutrition, health behaviors, and sleep.

Given that we find a considerable decline in cognitive performance among the elderly who obtain NRPS pension benefits, it is worth placing our results in the context of previous empirical research. Using data from high-income countries (in the E.U. and the U.S.), Rohwedder and Willis (2010) and Mazzonna and Peracchi (2012) examine the effect of early retirement on memory performance. Both studies find considerable evidence of harmful effects on cognitive performance associated with early retirement, a phenomenon Rohwedder and Willis (2010) refer to as *mental retirement*.[8,9] The effect size in Rohwedder and Willis (2010) is more than a standard deviation of the cognitive score for individuals in their sample. Compared to these two studies, our results also

[8] Using data from the US, England, Canada, and 11 European countries, Rohwedder and Willis (2010), Bonsang et al. (2012), and Adam et al. (2007) examine how retirement rates influence cognitive functioning and find a significant negative effect between retirement and cognitive functioning. Conversely, Coe et al. (2012) find no conclusive evidence with data from the US Health and Retirement Survey (HRS). Other recent studies also examine the effect of the NRPS on other individual or household-related outcomes. Nikolov and Adelman (2009ab) examine the NRPS program's effects on intergenerational transfers and health behaviors.

[9] Using data from the U.S., Jones, and Yilmazer (2018) show a positive relationship between positive shocks to lifetime income due to variation in EITC income benefits and cognition among a sample of aging adults.

demonstrate adverse effects of early retirement on cognitive ability; however, our estimates are considerably lower than the estimates by Rohwedder and Willis (2010) and Mazzonna and Peracchi (2012).[10]

Furthermore, we formally test for a difference in program impact on cognitive performance by gender. Although we report a faster and more substantial cognitive decline among female beneficiaries of the NRPS[11], our analysis cannot reject that impacts are similar for men and women at the conventional level of statistical significance.

We contribute to the existing literature as we are one of the first studies to examine how access to a retirement pension plan affects cognitive performance in the context of a developing country, and our study relies on a rich new dataset supplemented by analyses of administrative records.[12,13] These datasets allow us to consider broader measures of cognition, and we shed light on possible mechanisms. Studying how retirement policies lead to enormous depreciation of human capital is especially important in China because of its population size and the growing share of its elderly population.[14] Second, we illuminate how program participation affects a

[10] Mazzonna and Peracchi (2012) find a negative effect of retirement on orientation, immediate recall, and numeracy skills. The effect size is approximately between 0.2 to 0.3 standard deviations of the raw baseline cognitive performance measures (considerably larger than the estimated effect sizes from our analysis). Other studies explore the link between retirement and cognitive decline with data from high-income counties (Adam et al. 2007, Rohwedder and Willis 2010, Bonsang et al. 2012, Coe et al. 2012, Mazzonna and Peracchi 2012, Bingley and Martinello 2013, de Grip et al. 2015). Additionally, we find that pension benefits provision leads to a more substantial impact on delayed recall than on other cognition measures. The "delayed recall" test is one of the most sensitive tests to distinguish the effects of normal aging from Alzheimer's disease (Laakso et al. 2000). Another distinct feature of our sample, which is likely a key driver of some of the differences across these studies, is that we rely on data from a rural sample in a developing country, whereas Willis (2010) and Mazzonna and Peracchi (2012) use data from high-income countries.

[11] If gender differences in cognitive decline exist, this fact could have alarming implications for pension policy. The average performance on cognition tests for Chinese females is much lower than Chinese males' performance; the gender difference is pronounced among older Chinese cohorts. Coupled with the fact that females have a longer life expectancy, a faster cognitive decline due to an earlier onset of retirement could be an additional contributor to a gender-based expansion of morbidity in older age.

[12] Recent studies examine the effect of retirement policies on health behaviors in the context of high-income countries (Eibich 2015) or developing economies (Nikolov and Adelman 2019b). Nikolov and Adelman (2019b) show that older adults with access to the NRPS pension program experienced significant improvements in several health measures, including mobility, self-care, usual activities, and vision. This study shows that the NRPS has a considerable negative effect on cognitive ability among the elderly. Therefore, it is important to underscore a potential explanation between the observed difference in impacts on cognition and proxies of health. The factors determining cognitive depreciation are likely different from the factors (inputs) into the health production function. We examine this issue, in Section IV, with analyses on potential mechanisms driving the cognitive decline among NRPS beneficiaries.

[13] Using a fixed-effects estimation, Cheng et al. (2018) examine the health implications of the NRPS using data from the Chinese Longitudinal Healthy Longevity Survey one year after introducing the NRPS in 2009. Although Cheng et al. (2018) use only one year of survey data after the NRPS introduction and they do not directly observe NRPS participation (Cheng et al. 2018, pp.57), there is an overlap between the health inputs reported in Cheng et al. (2018) and the potential inputs in the cognitive depreciation process. Therefore, we return to the issue of examining the potential mechanisms underlying changes in cognitive depreciation, due to the NRPS, among the elderly in Section V.

[14] China's population is aging rapidly. In 2007, approximately 11 percent of China's population was ages 60 and over, making up 21 percent of the world's elderly population (UN 2019). Our analysis focuses on China, the country with the largest population globally, home to 1.4 billion people.

broader set of cognitive domains than has been previously considered. Although some cognitive decline appears to be an inevitable byproduct of aging, faster onset of cognitive decline can have profound adverse consequences on various aspects of life in old age—for example, financial planning for retirement (Banks and Oldfield 2007) and medical treatment adherence to planning for sequential activities (Fillenbaum et al. 1988). This paper focuses on various proxies of cognition, such as episodic memory, which neurobiology research shows to be particularly sensitive to the aging process. Several studies highlight that this domain is the first to decline with aging (Souchay et al. 2000; Prull et al. 2000). The second reason relates to its provision of higher individual variation than other cognitive measures.[15] Third, this study uses data from the CHARLS, a survey harmonized with the U.S. Health and Retirement Study (HRS) and other sister health surveys in high- and middle-income countries.[16] The survey harmonization of cognition measures across surveys can enable additional international comparisons of retirement-related patterns human capital depreciation related to cognition across countries. Finally, we provide suggestive evidence on the underlying mechanisms leading to cognitive decline among the elderly.

Therefore, the implications of this study's findings are likely to affect a significant portion of the global population, which additionally underscores the importance of the findings from a welfare standpoint. The study setting is unique because we analyze data from China's rural areas, whose demographic and economic activity resembles low-income countries. Therefore, our findings have important implications for other low-income countries.

[15] For example, the word learning and recall tasks do not exhibit floor or ceiling effects (excess of maximum or minimum values). The individual distribution of the scores does not exhibit extreme observation bunching around minimum and maximum values. Related to this, the CHARLS includes several cognitive measurements. We combine this into an aggregate cognition index. The use of an index of outcomes, a method based on the approach adopted by Kling, Liebman, and Katz (2007), addresses the possibility that the results are an artifact of multiple hypothesis testings and provides robust evidence of the global impact of the program.

[16] Started in 1992, the Health and Retirement Study is a biennial longitudinal survey. The main objective of the survey is to facilitate the interdisciplinary study of aging and retirement. The survey's core component collects data on a wide array of topics, including current health, cognition, current labor market participation, employment history, and subjective expectations about future events. Over the last three decades, it has collected information on more than 43,000 individuals in the U.S. The HRS has been a model worldwide for similar surveys that specifically examine health and retirement issues. Currently, harmonized constructs on health and demographic information exist across 18 longitudinal aging sister studies (e.g., Ageing and Retirement in Europe, the English Longitudinal Study of Ageing, Longitudinal Aging Study in India, Health and Aging in Africa: A Longitudinal Study of an INDEPTH Community in South Africa, Costa Rican Longevity and Health Aging Study) around the world. Because data in these surveys is calibrated based on the U.S. HRS, they allow for analysis of data that is harmonized for cross-national comparisons. More information on this data project is available at https://hrs.isr.umich.edu/about/international-sister-studies.

[16] The international family of HRS studies has adapted the HRS cognition measures across the HRS sister surveys (Langa et al. 2020), including the CHARLS (Meng et al. 2019) that, with caution, allow integrated analysis (e.g., Rohwedder and Willis 2010).

The remainder of the paper is organized as follows. Section II outlines the implementation of the rural pension scheme and summarizes the data. Section III presents the identification strategy. Section IV presents the results. Section V reports additional robustness checks and bolsters the validity of the empirical approach. Section VI concludes.

II. Background and Data

A. China's New Rural Pension Scheme

History and Expansion. Before the 1980s, China's public policies regarding its elderly population were mostly decentralized. Although some pension programs existed, they were initiated on an ad hoc basis and financed at the provincial level.[17] Weaknesses in the old-age pension system established in the 1950s began to surface in the early 1980s as the country moved more aggressively towards market reforms and a market-oriented pension system.

Besides the decentralized pension system's crumbling fiscal sustainability, the country's demographics changed rapidly in the 1990s. In 1979, China introduced its one-child policy to meet a population target of 1.2 billion by 2000. The government also expected zero population growth by 2000, and its targeted growth rate for the 1980s was between 0.5 and 1 percent. The combination of considerably lower fertility due to the one-child policy and reduced mortality caused the population structure to shift towards older age groups, resulting in a rapidly aging population.[18] Chinese government policy was necessary to tackle the growing demographic

[17] Vilela (2013) reviews the history and the evolution of China's pension policy since establishing the new People's Republic of China in 1949 up to 2013. The study posits that the country's policies toward its old-age segment have been moving away from its historical focus on formal-sector workers to emphasize universal coverage of formal and informal workers alike. Vilela (2013) highlights three distinct historical phases of the country concerning retirement policies: the "Iron Rice Bowl" (1949–1978), formal-sector pension reform and rural pension piloting (1978–2001), and a gear change in pension expansion (2003 to the present).

[18] The population growth rate in most age groups remained stable in the period from 1950 to 1980. This pattern produced an expansion of the age pyramid and resulted in the relative stability of the population's age-sex distribution. However, from 1964 to 1982, the oldest age groups experienced a considerable proportional increase, while the percentage distribution of the two youngest age groups declined substantially.

challenges and old-age poverty in rural areas. The government introduced a rural pension program in 1986. However, a combination of poor governance and financial challenges halted this expansion during the 1990s.[19]

In the early 2000s, the rural pension system faced continuing challenges related to financial sustainability. Under the new Hu-Wen administration that assumed leadership in 2003, the government ostensibly adopted a reform-oriented social protection system approach. Based on a 2008 pilot project initiated in the city of Baoji in the northwestern province of *Shaanx*, the administration proposed an ambitious transformation of its pension system.

In 2009, China launched the NRPS. Participation in the NRPS was available to all rural residents over the age of 16, provided they had not already enrolled in an urban pension scheme. The rollout of the program occurred based on administrative areas called *Hukou,* a system of household registration. Participation in the new program was voluntary, and individuals who were 16 years or older could contribute towards benefits that they could receive once they reach the age of 60. The rural program extended grandfathering conditions for residents who had already reached age 60 when the program launched. These individuals were eligible to receive a basic monthly benefit of 55 RMB, provided they had children who made monthly contributions towards the program.[20] Participants between the ages of 45 and 60, with fewer than 15 years of

[19] Financed by individual voluntary contributions and matching funds from local governments, the program covered state enterprise employees and individuals previously covered by the Basic Old-Age Insurance Scheme, a program mainly designed for urban employees (Liu and Sun 2016). Under the new system in the 1980s, the pension scheme introduced coverage quotas in urban and rural systems. Following a decade of pension reforms throughout the country, the Third Plenary Session of the 14th Communist Party Central Committee in 1994 additionally set targets for expanding the existing old-age social insurance system. The framework adopted by the party called for a multi-pillared system combining a fundamental social pillar with supplemental enterprise-sponsored pensions and individual savings for old age. By 1998, the pension system covered two-thirds of rural counties or 2,123 counties within 31 provinces. However, a combination of poor governance and additional financial challenges, complicated by the Asian financial crisis in 1997, halted the rural pension program's expansion, and it was substantially scaled back by the early years of the 2000s. Under the system, pension coverage declined from 80.25 million participants in 1998 (approximately 11 percent of the total rural population) to mid-50 million in 2007.

[20] The central government fully subsidizes the basic pension in Central and Western provinces and splits the cost with local governments in Eastern provinces (Cai et al. 2012).

contributions, had higher monthly contributions to make up for the delayed onset of participation before age 45.

The NRPS aimed to achieve full geographic coverage in rural areas by 2020 (Dorfman et al. 2013; Cai, Giles, O'Keefe, and Wang 2012). The program covered 23 percent of districts (or 29 million beneficiaries) by the end of 2010, and over 60 percent of districts (or 134 million) by early 2012. Program coverage expanded between 2009 and 2013 (depicted in Figure 1).[21] By the end of 2011, over 50 percent of rural residents contributed to the NRPS. Total participation in the program grew from 87 million to 326 million people from 2009 to the end of 2011 (Quan 2012).[22]

[Figure 1 about here]

Three major factors account for the NRPS expansion between 2009 to 2011. First, the high economic growth rate played a considerable role. Between 2009 and 2011, the economy grew at an average annual rate of 9.3 percent, which provided robust fiscal capacity for the massive social protection program rollout. Second, because of increasing income inequality and demographic pressures, demand for the program was substantial. Third, pension reform and the program expansion into the rural areas were fundamental political priorities for the Hu-Wen administration.

Program Eligibility and Benefits. Each individual who contributes to the pension program is entitled to program benefits comprising two components: (1) basic pension benefits of at least 55 RMB a month, and (2) individual account funds based on individual contributions and government subsidies. Regarding the individual account funds, an individual can typically

[21] Appendix Figure A1 shows the fraction of the elderly population (within a community) that received NRPS benefits over the four-year time span.

[22] We examine whether communities that implemented the NRPS earlier differed from the age characteristics of communities that adopted the program later. We find no evidence of differences in the average age. However, areas that implemented the NRPS versus non-participating areas could differ on other socio-economic characteristics, which we address in Section III.

opt for one of five annual contribution levels: 100, 200, 300, 400, or 500 RMB. These contribution levels are approximately equivalent to two to eight percent of China's 2009 rural annual per capita net income. Once a person reaches age 60, the total amount available for payouts is based on prior contributions and matched by local government funds. Local governments must match at least 30 RMB annually per individual contribution.[23] Based on the chosen contribution level, a government subsidy is added to each individuals account (e.g., a government subsidy of 30 RMB/year for a contribution level of 100 or 200 RMB/year; a government subsidy of 40 RMB/year for a contribution level of 300 RMB/year; a government subsidy of 45 RMB/year for a contribution level of 400 RMB/year). In this way, the current NRPS design concentrates incentives on the ex-post subsidy (the financing of the pension benefit) and has the advantage of simplicity. Individual contributions and government subsidies are deposited into the beneficiary's account. According to the People's Bank of China (the Central Bank of China), the interest rate is the one-year base rate, which was approximately 2.5 percent in 2011. When the central bank changes the base rate, the pension plan's interest rate adjusts accordingly, and interest compounds annually.

Based on data collected from early program implementation, nearly 50 percent of participants opted for the minimum annual contribution of 100 RMB (Dorfman et al. 2013). Individual contributions are voluntary, and they range from 100 to 500 RMB (two to eight percent, respectively, of the average rural wage in China) on an annual basis. [24,25]The pension

[23] This amount is independent of the individual contribution amount and may be subject to higher match amounts depending on the local government's budget. This match amount is less than one-to-one, given the minimum contribution is 100 RMB and the basic match is 30 RMB. Lei, Zhang, and Zhao (2013) show that most program participants, as of 2012, contributed 30 RMB per person per year.

[24] A participant may stop contributing for a few years and make up for the missed contributions later; they would only lose the subsidies for years when they did not contribute. Partial withdrawal from the accounts is not allowed. Participants can withdraw all of their savings under the following conditions: migration, change from a rural *hukou* to an urban *hukou*, or enrollment in an urban pension plan.

[25] Lei, Zhang, and Zhao (2013) conduct various simulations on the present value, factoring in the opportunity cost of accumulated pension accounts and the present value of the accumulated benefit using the current (at the time of the study) rate of return for the NRPS program. They show that the best investment strategy – if individuals maximize the net present value of the NRPS contributions – is choosing the lowest premium level and contributing as late as possible. They show that when the subsidy is 30 RMB/year, the optimal option is to contribute less than 21 years and choose the lowest premium level (100 RMB). This strategy has a positive net benefit (under additional assumptions about the annual interest rate, the

program also provided a fixed monetary pension payment. However, contributors need to contribute at least 100 RMB per annum to satisfy the vesting requirements for the basic benefit pension.

Individuals aged 60 or older can start receiving the basic pension every month without making any contributions if all their eligible children living in the same village participate in the NRPS. Individuals between the ages of 45 and 60 are eligible to receive the basic pension after age 60 if they contribute each year until they reach 60. Those under age 45 are eligible to receive the basic pension after age 60 if they contribute each year for at least 15 years.[26] The pension payouts do not depend on an earnings test, and therefore, participants can continue to work if they wish to do so when they start receiving their pension income.

The monetary benefits paid out to participants follow the "*139 Rule*", based on the average life expectancy (in months) at age 60.[27] The rule follows a basic formula for calculating the monthly payment: it takes the accumulated balance in the individual account and divides it by 139. Thus, the monthly payment comprises the basic pension plus the individual account balance divided by 139.

B. Data

Our primary analysis draws on data from the CHARLS, a nationally representative survey with rich data on various cognition proxies. We compare the cognitive outcomes of individuals with access to NRPS program benefits. The survey also collects data on the NRPS,

timing when the pension benefits are claimed, and the annuity factor). The return on contributions above the lowest contribution levels is limited to the return on pension assets (i.e., the one-year bank deposit rate). Therefore, participants have a strong incentive to opt for the lowest contribution level (100 RMB). The matching government subsidy of 30 RMB for an annual contribution of 100 RMB is likely too low to incentivize workers to contribute beyond the 15-year vesting period for the basic benefit.

[26] Except for the local government subsidy, the individual account is completely inheritable upon the recipient's death. Individual account balances are not forfeited upon death. If pensioners die sooner than 139 months after age 60, their heirs receive a lump sum payment for the amount of the remaining balance in the individual account minus any government subsidies. If pensioners live more than 139 months after age 60, they still receive a monthly pension as an annuity until death.

[27] The individual account has a rate of return equal to the People's Bank of China's one-year deposit rate. The "139" Rule was adopted based on the already established Urban Pension Scheme.

including individual participation in the program. The second part of our analysis draws on the China Health and Nutrition Survey (CHNS). We use data from a second source survey because the CHNS collects information on years before the NRPS. The survey data also enables us to conduct additional robustness checks, described in Section V.

The China Health and Retirement Longitudinal Studies (CHARLS). The CHARLS is a nationally representative survey that collects information on households that comprise at least one person who is 45 years or older.[28] The sampling frame comprised all Chinese provinces and counties, except for the Tibet province. The CHARLS collects data on demographic characteristics, family structure, cognition, health, pension and retirement, work, household wealth, income, and consumption. The survey's data collection started a year after the launch of the NRPS.

Our analysis sample consists of 15,990 individuals across two waves from 429 communities in 121 cities across 28 provinces based on an individual-level panel. The raw sample totals 17,708 individuals living in 10,287 households in 450 villages/urban communities in 150 cities/districts across 28 of China's 30 provinces, excluding Tibet. The 2011 baseline wave interviewed 10,257 households with 18,245 respondents aged 45 and over.[29] The follow-up 2013 wave covered 10,979 households (or 19,666 respondents).[30] The CHARLS directly collects information on individual participation in various government programs, including the NRPS. We drop observations with an urban Hukou status for our primary analysis sample because individuals attached to an urban *Hukou* are ineligible to participate in the NRPS but instead

[28] For context, the life expectancy at birth was 74.5 years for men and 76.7 years for females in 2018. (UN 2019). The average retirement age in rural areas is 61 years (SOA 2016). The average household income (excluding pension income) was 26,022 RMB; the median household income was 11,680 RMB. On average, NRPS beneficiaries received 233 RMB/year (i.e., approximately 1 percent of the average income or 2 percent of the median income).

[29] Figure B1 shows the geographic coverage map for the CHARLS survey. Initially, 19,081 households were sampled, where 12,740 had age-eligible members, of which 10,257 responded.

[30] The interviewers followed up with 88.6 percent of the original respondents and 89.6 percent of original households. The 2013 Wave added 2,053 new households comprising 3,507 individuals.

participate in urban pension schemes.[31] Furthermore, our analysis does not include individuals over the age of 60 with no children at the time of program coverage because they were ineligible to participate in the NRPS. We can directly observe NRPS participants and non-participants.

Proxy Measures of Cognition. A second attractive feature of the CHARLS is that it directly tests cognition based on several proxy measures based on comprehensive research on aging and cognition, and measures used in the HRS (Ofstedal 2005).[32] The first cognition measure tests episodic memory captured via verbal learning and several recall tasks.[33] The second cognition measure tests mental intactness.[34] Although we analyze all cognition measures tested in the survey, we pay special attention to the episodic memory domain for two reasons. First, several studies highlight that this domain is the first to decline with aging (Souchay et al. 2000; Prull et al. 2000).[35] In addition to several cognitive tests, respondents are asked to rate their memory based on a 5-point scale.[36] Based on this scale, we create a binary indicator to denote if an individual is in good physical health.

[31] The Urban Social Pension Scheme was established in 2011 and rapidly expanded to cities with robust fiscal capacity. The program is voluntary and is offered to urban residents aged 16 and over who are not employed in the formal sector. The program features a two-tier system, which consists of a pay-as-you-go social pooling component and individually funded accounts.

[32] The HRS cognition measures, and the ones used in the CHARLS, accounted for several essential considerations. First, the measures represent the major dimensions of cognitive functioning and can differentiate across a range of cognitive abilities. Second, the measures can identify respondents who exhibit some form of cognitive impairment. This second consideration guided the inclusion of a traditional mental status measure that can differentiate individuals at the low functioning end of cognitive abilities. A third consideration included screening for early signs of dementia, or in the case of onset, for its subsequent progression.

[33] CHARLS uses the HRS version of the CERAD immediate and delayed word recall to measure episodic memory (Ofstedal et al. 2005). *Episodic memory* is a necessary component of reasoning in many dimensions. The two tasks that capture verbal learning and recall are immediate and delayed recall. After approximately four minutes after other questions, the respondent is asked again to recall the nouns, without reading the words a second time. Word recall tests are collected to assess individuals' short-term and long-term cognitive impairment. For the immediate recall test, surveyors randomly assign respondents with a list containing ten common words. The respondent is given two minutes to recall as many words as he/she can remember. The immediate recall score ranges from zero to ten and provides the number of words recalled correctly. Following this recall, the respondent answered unrelated questions for several minutes until prompted to recall the original word list. This procedure captures the delayed recall score, which ranges from zero to ten.

[34] This mental intactness task comprises recognition of date: (month, day, year, season, day of the week), self-rated memory (excellent, very good, good, fair, and poor), and serial subtraction of 7s from 100 (up to five times). The respondent is also asked to redraw a picture of overlapping pentagons. We compute the sum of two scores based on these measures—the immediate and delayed recall—for a total word recall score, ranging from 0 to 20. Low scores on this total word sum are indicative of low memory capacity and short storage duration.

[35] For example, the word learning and recall tasks do not exhibit floor or ceiling effects (excess of maximum or minimum values). The individual distribution of the score does not exhibit extreme observation bunching around minimum and maximum values. The CHARLS collects additional cognitive measures elicited by the survey respondent.

[36] The 5-point scale used for the measurement of the episodic memory is as follows: (1) Excellent (2) Very Good (3) Good (4) Fair (5) Poor.

Cognitive ability comprises two components: fluid and crystallized intelligence (Brown 2016). Fluid intelligence is the ability for abstract reasoning, memory recall, and drawing inferences. Thinking on the go and the act of solving novel problems involves fluid intelligence. Crystallized intelligence, on the other hand, is the accumulated stock of knowledge. The knowledge people obtain through schooling, and life experience is crystallized knowledge; this knowledge builds up into human capital stock. Our first measure, episodic memory, captures aspects of fluid intelligence as it encompasses the ability to reason and recall information from memories (McArdle et al. 2011). In contrast, the mental intactness measure captures both fluid and crystallized intelligence as the measure pertains to the ability to infer and access the stockpile of knowledge, referred to as human capital (McArdle et al. 2011).

We combine data from the following factors: perceived memory status (subjective status), knowing the current month (orientation), serial-7 score (working memory), immediate recall score (memory capacity), and delayed recall score (memory duration). Using principal component analysis (PCA), we reduce these multiple measures into one composite index.[37,38,39]

We report the summary statistics of our sample in Table 1. Among the sample of participants and non-participants, 70 percent and 69 percent were employed in the baseline, respectively. Seventy-two percent of participants and non-participants work in agriculture. The rural sample reported low educational attainment levels—approximately 46 percent to 48 percent report having completed at least a secondary school degree. In terms of health, 27 percent of participants and 26 percent of non-participants report being in "poor/fair" health status.

[37] This index provides a normalization of cognitive memory status, where negative (or low) values are associated with poor memory functioning. This index is an overall cognition proxy in the analyses that we present in the subsequent section

[38] We use a PCA method to transform the set of proxy variables for cognition into an aggregate index. We do so by first standardizing each cognition proxy. We then compute the covariance matrix for all cognition measures. Third, we compute the eigenvectors; combined, they contain the same information as the original variables. The first component, based on the largest eigenvalue, contains the most information by design, whereas the last component contains the least. We reduce the original cognition proxies into one index by retaining the component with the largest variance (eigenvalue). The overall PCA index, Cognitive Memory Index, has a mean of zero and a standard deviation of 1.42.

[39] Online Appendix B, Table B2 reports the index component loadings based on the survey's cognitive measures.

[Table 1 about here]

Regarding the variables of interest in the study, survey participants report a low average on various cognitive measures in the baseline period. Program participants and non-participants reported their memory as being "at least good" 15 percent and 18 percent, respectively. [40]

China Health and Nutrition Survey. As the CHARLS does not collect cognition data before the start of the NRPS, we rely on data from the China Health and Nutrition Survey (CHNS), a survey conducted by the University of North Carolina at Chapel Hill which covers 1989 to 2011, a period overlapping with the start of NRPS. The CHNS covers approximately 19,000 individuals in 15 provinces spanning 216 primary sampling units (PSUs).[41] The survey aimed to collect data on various economic and social determinants of individual health and nutritional status (UNC-Chapel Hill 2010). Adults ages 55 and older recorded their daily living activities and performed various cognition tests. The CHNS also collected information on proxy measures of memory and cognition, similar to those collected by the CHARLS. The CHNS adopted similar cognitive screening items because because it adopted measured adapted from the U.S. HRS survey. The same cognitive screening test was used in the three waves of the CHNS among adults aged at least 55 years.[42,43]The CHNS sampling areas overlap with the ones sampled by the CHARLS.[44]

[40] Participants scored slightly higher on the word recall tests. The average score on the immediate word recall task for participants was 3.93 out of 10 (non-participants average 3.77 out of 10). Similarly, the delayed recall score was higher for participants than for non-participants, 2.91, and 2.89. Approximately 84 percent of participants and non-participants correctly named the current month. Based on the PCA, the cognitive memory index exhibits a higher average for participants than for non-participants, 0.06 and 0.00, respectively.

[41] The survey covered the following provinces (see Figure 2): Beijing, Chongqing, Guangxi, Heilongjiang, Henan, Hubei, Hunan, Jiangsu, Liaoning, Shaanxi, Shandong, Shanghai, Yunnan, and Zhejiang. The CHNS collected data on fourteen provinces in contrast to the twenty-eight provinces in the CHARLS.

[42] The cognition tests tested immediate and delayed recall of a two-word list, counting backward from 20, serial-7 subtraction, and memory orientation. The scores for immediate and delayed recall ranged from zero to ten. Counting backward and serial 7s were used to assess attention and calculation, with scores ranging from zero to seven. Orientation was assessed by asking the participant the current date (one point each for a correct response on the year, month, and date) and the name of the tool usually used to cut paper (one point). Higher scores on all items suggest improved cognitive performance.

[43] We use this survey for analysis on pre-trends because there are no data in CHARLS for the outcomes we analyze prior to the baseline period, which is 2011.

[44] Figure B1 depicts the geographic coverage of the CHNS. Appendix Table B1 compares the summary statistics for the CHARLS and CHNS for adults ages 45 and over and living in rural areas. CHARLS only collects information on adults ages 45 and over.

III. Empirical Strategy

Our primary identification strategy relies on variation across municipalities in the NRPS implementation. We exploit a source of identifying variation due to the program's staggered rollout between 2011 and 2013 across its rural parts.[45] The DDD analysis exploits the rollout at the community (*shequ*) level, an administrative level within a county that encompasses several neighborhoods. For each community, the CHARLS administered a community questionnaire that collected data on its natural environment, employment, financial status, and social protection program coverage. Our main objective is to examine how the rollout of the program affected individual cognition among eligible individuals.

A. Estimating Equations and Triple Difference Estimation

Given the NRPS's staggered rollout at the community level, our identification strategy relies on when communities adopted the pension scheme.[46] Using variation at the community level (some communities implemented the program, and some did not), the timing (some areas adopted the program earlier than others did), and the eligibility for pension benefits, we use a DDD estimator to estimate the effect of the NRPS on cognition. Although the identifying variation comes from areas (i.e., communities) treated between 2011 and 2013, we perform our analysis at the individual level.

Based on information from the CHARLS, we construct a variable, $OfferNRPS_{ct}$, which indicates the participation status (whether a community c implemented the NRPS program at time t). Data from the CHARLS records the residence location, whether the local municipality

[45] Figure B2 reports the distribution of individual NRPS participation, at the community level.
[46] Demographic information in the CHARLS is only available at the community (*shequ*) level.

(*shequ*) implemented the NRPS, and whether the respondent enrolled in the NRPS.[47] This process allows us to define the variable $OfferNRPS_{ct}$ based on responses from the individual-level data.[48] We examine the impact of the NRPS provision on cognition using the following specification[49]:

$$(1) \qquad Y_{ict} = \beta_0 + \beta_1(OfferNRPS_{ct} \times Above60_{ict}) + \beta_2 Above60_{ict} + \beta_3 \boldsymbol{X}_{ict} + \phi_c + \mu_t + \phi_c \times \mu_t + \varepsilon_{ict},$$

where Y_{ict} is the cognition outcome and $Above60_{ict}$ is equal to 1 if the respondent is aged 60 and over.

The coefficient of interest in (1) is β_1. It captures the intent-to-treat (ITT) estimate of the average effect of the NRPS program on the average outcomes of eligible individuals aged 60 and over who live in a treated community, regardless of whether the individual decides to participate in the program. $\boldsymbol{X}_{ict}$, is a vector of individual-level controls; ϕ_c and μ_t are community-level and time fixed effects. Community-level fixed effects control for time-invariant characteristics that affect the likelihood of a community implementing the NRPS.[50,51] Community-time fixed effects, $\phi_c \times \mu_t$, control for community differences during the NRPS implementation. We estimate specification (1) with and without individual fixed effects. The DDD design is the most appropriate choice, as it controls for potential region-specific effects and is based on a similar

[47] It is not common for individuals to stop participation in the NRPS once they enrolled in the program; however, 705 individuals indicated they were enrolled in 2011 but not in 2013. At the community level, eight communities (out of 350 communities used in our sample) implemented NRPS in 2011 but not in 2013. Only 7.5 percent of the 705 individuals who stopped participating in the NRPS resided in the eight communities that indicate terminating their NRPS implementation. Therefore, these individuals are not a threat to the validity of our community-level instrumental variable.

[48] If no individuals indicate having NRPS at time *t* in community *c*, then $OfferNRPS_{ct}$ equals 0. If at least one person reports participating in the NRPS, $OfferNRPS_{ct}$ is set to 1. We address potential concerns regarding measurement error and associated bias in the estimated coefficients based on this approach with additional robustness checks that we present in Section V.

[49] Specification (1) is estimated using a three-way error component model (community, year, and community-year fixed effects) as previously used in the applied econometrics literature (e.g., Andrews et al. 2006).

[50] In In several robustness exercises, we also include a richer set of age-related controls (e.g., age, age-squared), and results remain robust in these specifications. We cluster the standard errors by community and age groups.

[51] In Online Appendix B, we report additional robustness checks where we cluster the standard errors by community and age. Our results are robust to community and age-specific clusters.

policy rollout in other empirical studies, such as Katz (1996) and Gruber (1994). We opt for an estimation approach based on the DDD method instead of the DD approach for two primary reasons. The DDD estimation accounts for two kinds of potentially confounding trends: changes in the cognitive performance of individuals 60 in the communities that implemented the NRPS and changes in the cognitive performance of people living in all communities that implemented the NRPS (possibly due to other local policies that could affect cognitive performance).

If the variation in program implementation across communities is unrelated to other community-related shocks, specification (1) will produce an unbiased estimate of β_1, the coefficient of interest. The identification assumption, for the DDD design, requires that the relative outcomes (cognitive performance) of people over the age of 60 and people less than 60 in the areas that implemented the NRPS trend in the same way as the relative outcome of people over age 60 and people less than age 60 in the areas that did not implement the NRPS, in the absence of the NRPS (Olden and Moen, 2020). To check whether the triple difference is an appropriate strategy to examine the NRPS's effect, we test the common trends assumption for the pre-policy survey data based on the empirical approach in Autor (2003). We examine the trends of various cognition measures between treated and non-treated areas *before* the NRPS program launch in 2009. Since all survey data from the CHARLS is collected post-NRPS program, we analyze data on the pre-trends of our study outcomes for the three CHNS waves that collected cognition measures.

The primary data challenge for the analysis using the CHNS is that the community identifiers or geographic-level variables do not match in a one-to-one fashion with the CHARLS. Therefore, only for this empirical exercise using the CHNS, we redefine “treated” and “control” units at the province level (as opposed to the community *shequ* level) to rely on the geographic

variables available in the CHNS. Once we reconstruct the analysis on the province level, we test the common trends assumption at the province level using data from the CHNS from 2000 to 2009. Furthermore, and only for this empirical test, we underscore that the treatment status definition for a province in the CHNS data for 2004 to 2009 hinges on two critical features. First, it relies on the baseline data from the CHARLS. Second, the definition is based on the percentage, within a province, of communities that report NRPS implementation. Therefore, our definition of a "treated" province relies on the percentage of communities within a province that indicate (based on survey data from the CHARLS) that they participate in the NRPS program. In other words, we define a province as treated based on a continuous (treatment intensity) variable for a given province. Based on this treatment definition, we code the treatment status for each province as a binary variable. The treatment status is set to one if more than a given threshold of communities reported participating in the NRPS, and zero otherwise. To define a "treated" province, we choose a threshold based on the percentage of communities (within a province) that indicate that they participate in the NRPS.[52,53] Based on this reconstructed definition of "treatment" and only for this formal test of the *common trends'* assumption, we use data from the CHNS before the NRPS was launched. Our analysis for this test uses data on cognition outcomes from the CHNS that mirror the cognition proxies collected by the CHARLS. Using the CHNS data, before 2011, on cognition measures from 2000, 2004, and 2006 waves, we can estimate the following specification:

$$Y_{ipt} = \beta_0 + \beta_{-3} D_{ipt-3} + \beta_{-1} D_{ipt-1} + \phi_p + \mu_t + \phi_p \times \mu_t + \varepsilon_{ipt} , \quad (2)$$

[52] We use a binary definition of treatment status for each province. We define a province as "treated" (=1) if more than 67 percent of all communities within this province implemented the NRPS, based on baseline data. In addition to using this threshold, we conduct additional sensitivity analyses based on alternative threshold choices. In additional sensitivity analyses, we vary the threshold choice to a lower (50 percent coverage rate) or higher value (70 percent coverage rate). Based on these alternative threshold choices, we redefine each province's treatment status and re-estimate our specifications.

[53] The CHNS does not sample from the same communities/villages as the CHARLS, so we rely on our definition of treated and control provinces based on the CHARLS to test data in the CHNS.

where Y_{ipt} is the cognition proxy, and ϕ_p, μ_t, and $\phi_p \times \mu_t$ are province, time and province-time fixed effects, respectively. Because of the triple-difference estimation, we include the triple interactions $D_{ipt} = (OfferNRPS_{pt} \times Above60_{ipt})$ for the first and last pre-treatment periods a.[54] D_{ipt} is defined in the same way as in our main triple-difference specification; subscript p denotes the province. The results reported in Online Appendix Table A1 provide clear evidence that β_{-3} and β_{-1} are insignificant. Based on this test, we fail to reject the hypothesis that trends in the treatment and control areas' outcomes are the same.[55] Therefore, this exercise provides no empirical evidence to undermine the validity of the common trends assumption.

B. Instrumental Variable Estimation

To address the possibility of endogenous individual participation, we augment the DDD analysis by instrumenting individual program participation with an instrument at the community level. We re-estimate specification (1). However, we also use $OfferNRPS_{ct}$ to instrument for individual participation in the NRPS, an instrumented difference-in-differences design as in Hudson, Hull, and Liebersohn (2017).[56] We code $OfferNRPS_{ct}$ as a binary variable, the variable is set to 0 if no individuals participate in the NRPS. It is set to 1 if the community witnesses at least 1 participant. We re-estimate the following specification:

$$(3) \qquad Y_{ict} = \beta_0 + \beta_1(\widehat{NRPS}_{ict} \times Above60_{ict}) + \beta_2 Above60_{ict} + \beta_3 \boldsymbol{X}_{ict} + \phi_c + \mu_t + \phi_c \times \mu_t + \varepsilon_{ict}.$$

$$(1) \qquad Y_{ict} = \beta_0 + \beta_1(OfferNRPS_{ct} \times Above60_{ict}) + \beta_2 Above60_{ict} + \beta_3 \boldsymbol{X}_{ict} + \phi_c + \mu_t + \phi_c \times \mu_t + \varepsilon_{ict},$$

[54] In this specification, the second pre-treatment period is omitted.
[55] This test is based on analysis at the province level, as described above. Therefore, the number of observations only in this sample is considerably lower than in the primary sample (at the community level). It is possible that the low number of observations could lead to an underpowered inference for this test. However, Table A1 also shows that the estimated coefficients are statistically insignificant and unstable across different threshold specifications, which further undermines any evidence of robust differences, at the province level, in pre-trends.
[56] Hudson, Hull, and Liebersohn (2017) note the identifying assumptions for the DDD-IV estimation: the typical exclusion restriction for an IV, parallel trends (growth paths of both treatment and outcomes need to be independent of the actual instrument assignment), monotonicity (the effect of the instrument, the NRPS availability, on the NRPS takeup, needs to be monotone).

$\widehat{NRPS_{ict}}$ represents individual enrollment in NRPS, and we instrument it with $OfferNRPS_{ct}$. $\boldsymbol{X}_{ict}$ is a vector of individual-level controls. ϕ_c, μ_t, and $\phi_c \times \mu_t$ are community-level, time, and community-time fixed effects, respectively.

IV. Main Results

A. Impacts on Cognition Measures

We start by examining the impacts on cognitive outcomes based on specifications (1) and (3). Table 2 reports the results. Columns 1 through 8 report the results for the various cognition proxies for the immediate recall measure, delayed recall, total recall, and memory index, respectively. These results are based on estimating specification (1); therefore, they are the intent-to-treat estimates on the effect of program availability in a community on the various cognition measures. The results in all columns provide striking evidence of negative cognitive impacts among individuals aged 60 and above who live in NRPS program areas. In Table 2, we report the results based on the 2SLS approach.

The results in both tables indicate a striking pattern of adverse effects on all cognition measures. On the immediate recall test, individuals in NRPS program areas aged 60 and above score, on average, worse by eight percent of one standard deviation. For the delayed recall test (with individual fixed effects), individuals in program areas score worse by approximately 14 percent of one standard deviation. Program availability also has a considerably negative effect on the cognitive index. The index combines the cognition measure on mental intactness, described in Section II. On average, the NRPS benefits provision leads to a 0.12-point reduction in the composite score (equivalent to about 9 percent of one standard deviation) for the intent-to-treat specifications. Based on the 2SLS estimation, the effect size associated with the NRPS is

doubled, as reported in Table 2 Panel B. When comparing the effect size estimates across all columns, the largest negative effect is on the delayed recall cognition measure. This effect size, associated with the delayed recall measure, is approximately double the effect size for the other two cognition measures. Neurological studies document that this specific proxy measure of cognition is a useful predictor of dementia in adulthood (Welsh et al. 1991, Laakso et al. 2000).

[Table 2 about here]

In addition to the impacts of access to the program regardless of program participation, Table 2 reports the treatment-on-treated estimates based on specification (3). The effect size estimates reported in Panel B are approximately double the effect size estimates based on the ITT specification.[57] The results reported in Panel B echo the pattern reported in Panel A—the effect of program participation on measured cognition is statistically significant and negative for all cognition measures.

The analysis so far has focused on data from the CHARLS, based on whether the community implemented the NRPS or the individual reported participating in the program. The CHARLS also collects data on actual retirement. However, it is essential to underscore that data on this variable is sparse. Using the formal definition of retirement and other employment-related variables available in the CHARLS, we reconstruct the definition of retirement.[58] Despite the lack of data (or possibly data quality issues) for this reconstructed variable, we report results based on specifications (1) and (3) using data on the self-reported retirement status. Table 3 reports the results.

[57] Table 2 reports the F-statistic associated with the first-stage estimation in the 2SLS specification. The F-statistic is considerably above the usual rule of thumb value of 10.

[58] This re-constructed binary definition of retirement is based on available data on any of the following CHARLS variables: the person completed retirement procedures in any survey wave, the reported number of days (or months, hours) worked is zero in three consecutive waves, the reported usual number of days (or months) per year is zero for three consecutive waves, the reported monetary retirement benefit is positive, the number of workdays missed for health reasons has been more than 300 per year for three consecutive survey waves, reported year of retirement is before 2009, and the survey respondent indicated that the formal retirement is processed.

[Table 3 about here]

Similar to the negative effect highlighted in Table 2, the reported results in Table 3 show that program participation leads to faster cognitive decline via the program's influence on retirement decisions. Although not all results pass the conventional levels of statistical significance, all effect size estimates are negative. It is essential to underscore that very few individuals respond to the retirement status question in this dataset. Program participants who retire exhibit decline (i.e., negative coefficients reported in Columns 1 through 6) in their performance on the cognition tests.

Next, we consider if the length of exposure to program benefits leads to a more substantial cognitive decline among beneficiaries. If the NRPS was the primary contributing factor for cognitive decline, the length of time one receives benefits should lead to a larger cognitive decline effect size. To examine this possibility, we categorize all program beneficiaries aged 60 or above into three categories: less than one year of exposure to NRPS benefits, between one to three years of exposure to NRPS benefits, and more than three years of exposure to NRPS benefits. Using these three groups, we estimate the effect of length of exposure to NRPS benefits based on the primary DDD estimator interacted with dummies for two of the categories related to the length of exposure (the reference group is less than one year of exposure to NRPS benefits). Table 4 reports the results of this estimation. We find that individuals who receive NRPS benefits for more than three years show a more substantial cognitive decline rate than individuals with shorter exposure duration. Therefore, we find empirical evidence consistent with the idea that increased duration of exposure to NRPS benefits leads to a more substantial cognitive decline among NRPS beneficiaries.[59]

[59] In analyses not reported here, we repeat the same exercise with a continuously defined variable capturing the length of exposure to NRPS benefits, and we find the same negative pattern.

[Table 4 about here]

B. Mechanisms

Obtaining access to the retirement program benefits likely changes a host of behavioral outcomes that influence human capital depreciation. To examine possible mechanisms leading to the faster cognitive decline among NRPS beneficiaries, we start by considering other retirement-related outcomes affected by access to the NRPS.

Specifically, we examine how NRPS participation affected four major groups of activities of program beneficiaries: (1) labor market activities, (2) mental stimulation, (3) social engagement, and (4) various health behaviors, time use, and health care utilization. Suppose we do not observe any change in an outcome that could play a mediating role in influencing cognition. In that case, we take this as an indication that the causal pathway does not operate via that mediating factor (or group of factors). Tables 5 and 6 report the results. The NRPS had protective effects on various health behaviors. Program participants reported a reduced incidence of regular alcohol drinking than in the previous year (Column 3 in Table 6). Program participants, on average, reduced the incidence of their cigarette smoking. NRPS participation had a positive effect on sleep patterns.

Conversely, NRPS participation led to a considerable adverse impact on the remaining categories: labor market activities, mental stimulation, and social engagement. Among NRPS participants who remained active in the labor force, we see a consistent pattern of decreased activity in the labor market (as reported in Columns 1 through 5, both for wage employment and non-agricultural self-employment activities). Furthermore, we see substantial evidence of decreased social engagement. NRPS participants report lower social community engagement or volunteering than individuals who did not obtain NRPS benefits. Furthermore, for many NRPS

participants, retirement leads to a less stimulating daily environment. We document suggestive evidence of decreased mental stimulation among NRPS participants, but the point estimates are not statistically significant. Lower social engagement, coupled with evidence of more sleep, is consistent with a story of an increased incidence of depression among NRPS participants. However, the CHARLS does not collect adequate data to address this particular channel.

[Tables 5 and 6 about here]

Furthermore, the results on potential channels warrant an explanation of the net effects on cognition. On the one hand, NRPS participation leads to numerous benefits on diet, smoking, and health behaviors. On the other hand, the program also leads to considerable adverse effects on other outcomes. Other channels of influence may also be at work. However, given the overall decline of cognition among program participants and this analysis on mediating mechanisms, it seems likely that the adverse program effects on mental and social engagement far outweigh the NRPS's protective benefits on various health behaviors.

C. Cognitive Decline: Heterogeneity by Gender

Next, we examine whether the NRPS impacts differ by gender. We address the issue by examining for heterogeneous effects, and we formally test for gender effects among program beneficiaries. We do this using a quadruple difference (DDDD) method estimating the following specification:

$$(1) \quad Y_{ict} = \beta_0 + \beta_1\left(OfferNRPS_{ct} \times Above60_{ict} \times Female_{ict}\right) + \beta_2\left(OfferNRPS_{ct} \times Female_{ict}\right) + \beta_3\left(OfferNRPS_{ct} \times Above60_{ict}\right) + \beta_4(Female_{ict} \times Above60_{ict}) + \beta_5(Above60_{ict}) + \beta_6(Female_{ict}) + \beta_7(\boldsymbol{X}_{ict}) + \phi_c + \mu_t + \phi_c \times \mu_t + \varepsilon_{ict}$$

Specifically, we focus on the estimate for β_1, which captures the NRPS treatment effects on cognitive decline among females. Table 7 reports the results for the heterogeneous treatment analysis. Panel A reports the intent-to-treat DDDD estimates, whereas Panel B provides estimates based on the 2SLS estimation approach.

[Table 7 about here]

The results reported in Table 7 echo the negative effect of the NRPS program on cognitive measures. Table 7 reports the results for individuals who live in areas that implemented the NRPS program (in Panel A) and the results for individuals who participated in the NRPS (in Panel B). When we compare the effect size estimates for the aggregated measures reported in Column 4, we note that the NRPS effect on females' cognitive decline is approximately double the effect among males. However, the negative point estimate for the DDDD estimator is not statistically significant.[60]

V. Robustness Checks

In this section, we perform a series of robustness exercises. First, we conduct a falsification exercise to test the validity of our empirical estimation. Specifically, we re-estimate our main empirical estimations with an individual sample based on individuals who cannot benefit from the NRPS. Therefore, estimating the main specifications using this sample of non-eligible individuals should yield non-significant results. Second, we address the possibility of measurement error related to either individuals misreporting their participation in the NRPS program or a community incorrectly indicating program implementation within its boundaries.

[60] We examine potential mechanisms by examining the effect of the NRPS on various mediating factors. We report the results in Appendix A (Table A2). The results suggestively indicate that the primary driver of the faster cognitive decline among women may be related to the labor market, particularly for females in self-employment.

The main results survive these extension exercises, which implies that our main results are unlikely to be driven by secular trends, alternative contemporaneous policies in the areas that implemented the NRPS, or unobserved shocks impacting the study outcomes.

A. Falsification Exercises

We now turn to several falsification exercises to bolster the validity of our estimated results. We construct a falsification exercise based on an alternative sample of people who are not eligible for the NRPS or its benefits. Therefore, in theory, when we rerun specifications (1) and (3), the coefficients of interest discussed in Section III for this alternative study sample should not be significant.

As we underscored in Section III, the NRPS program is only available to individuals who live in rural administrative districts, provided they are not enrolled in an urban pension scheme. In the main analysis and results presented in Section IV, we excluded urban pensioners and rural-residing elderly individuals without children because they are ineligible for the NRPS. However, for this falsification exercise, we reconstruct the analysis sample and employ the opposite approach. This sample comprises pensioners in an urban pension program or rural-residing elderly individuals (aged 60 and above) who did not contribute to the NRPS before reaching age 60 and have no children residing in rural administrative districts.

In this falsification exercise, we only perform the analysis based on a sample of individuals who: (1) live in rural areas but obtain benefits from an urban pension system; or (2) are elderly without children and who happen to live in rural areas. The main objective is to examine the primary estimation approach detects cognitive performance effects among individuals in this *placebo* sample. If specifications (1) and (3) yield no spurious results, this falsification exercise should produce non-significant estimates for the coefficients associated

with the NRPS effect on cognition outcomes.[61] As described in the main empirical approach, we re-estimate specifications (1) and (2) using data from the placebo sample. In Appendix A, Table A3, we report the results based on this falsification exercise.

Table A3 reports non-significant estimates for program impacts on cognition measures: immediate recall score, total recall score, and cognitive memory index. In other words, these results imply that urban pensioners who live in communities that offer the NRPS (relative to urban pensioners who live in communities that do not offer the NRPS) do not exhibit statistically significant differences in cognitive performance. The results based on this additional robustness check further bolster the validity of our main results presented in Tables 2 and 4; they are unlikely to be based on a spurious specification choice.[62]

B. Alternative Measures of NRPS Participation

We further explore the possibility that our primary analysis relies on either mismeasured individual participation in the NRPS or incorrect reports among municipalities (*shequs*) of NRPS program implementation. Either of these possibilities will yield measurement error in our program impacts and could produce biased impact estimates. Therefore, we perform additional consistency checks based on alternative approaches intended to measure NRPS participation.

Propensity Score Method Definition. First, survey responses based on the CHARLS may be incorrect, resulting in possible mismeasurement of actual NRPS participation by individuals

[61] This additional test assumes the absence of spillover effects between the group of individuals who are beneficiaries of the NRPS and the urban pensioners who live in the same communities that offer the NRPS. Still, they are not eligible (nor do they receive program benefits) for the NRPS. As we show in Section III, NRPS participants lower their social interactions. Therefore, if spillover effects occur via social interaction (a viable mechanism for social spillovers), then our analysis will pick up program impacts among non-beneficiaries who live in areas that offer NRPS, and we do not detect such changes.

[62] We also conduct an additional falsification test in which we re-estimate specifications (1) and (3) on a set of placebo outcomes. The selection of these placebo outcomes was based on no conceptual mechanism linking pension and program impacts. This additional falsification exercise was another attempt to examine the credibility of our main results. In Appendix Table A4, we report the results based on a set of four "placebo" outcomes. The four placebo outcomes are: a person's nationality being Han, the number of living daughters in the household, mother's educational level, and the number of living sons. Appendix A Table A4 reports the results. In Panel A, we report the ITT results, where Panel B reports the TOT results. In both panels, the results provide no empirical support of program effects on the set of placebo outcomes.

in our analysis. We address this possibility with an alternative measurement of individual NRPS participation status. To do so, we use data on personal characteristics, and we reconstruct the likelihood that a person participates in the NRPS. We redefine the NRPS participation status based on a propensity score matching approach. We predict the NRPS participation status (at baseline) based on a combination of individual characteristics, such as education, gender, parental education, and nationality. We use the baseline data for these variables. Using these characteristics, we then predict the propensity of NRPS participation, $\widehat{NRPS}_{ic}$, based on the propensity score matching method. The predicted participation, based on this estimation technique, is $\text{PrNRPS}_{\text{ic}}$. We construct an alternative measurement of the NRPS participation status variable by defining $PrNRPS_{ic} = 1$ if $\widehat{NRPS}_{ic}$ is greater than one standard deviation above the mean of $\widehat{\text{NRPS}}_{\text{ic}}$.

Next, we use this redefined measure of NRPS participation, and we re-estimate specifications (1) and (3). The NRPS program participation in this analysis uses the reconstructed variable ($PrNRPS_{ic}$), based on the estimation from the propensity score approach, as opposed to the estimation approach in Section IV (based on the self-reported variable in the CHARLS).[63] We report the results based on this alternative definition of program participation in Online Appendix Table B3. The reported results, using the alternative NRPS participation definition, bolster our main findings.

Community NRPS Participation Definition. We consider the possibility that there is a measurement error due to individuals misreporting their NRPS participation in the CHARLS. If true, this measurement error at the individual level could generate possible misclassification of

[63] Two key assumptions underlie this propensity score approach. First, the approach assumes that only observable (and time-invariant) characteristics determine selection into participation in the NRPS. Second, the method relies on the assumption that, in the absence of the NRPS, the age trend in cognitive functioning is the same between covariates, not used in our propensity score determination.

areas reporting they implemented the NRPS, an issue affecting our analysis based on specification (3).[64]

Therefore, we verify the robustness of our approach with an alternative definition of the variable $OfferNRPS_{ct}$. This additional exercise aims to correct possible contamination of what areas are *treated* (i.e., indicate NRPS implementation). We re-estimate specifications (1) and (2) but rely on a higher threshold that defines when the variable $OfferNRPS_{ct}$ (the variable that indicates community participation in the NRPS) switches from zero to one. Instead of relying on a threshold of at least one individual reporting NRPS participation to set $OfferNRPS_{ct}$= 1, we now use an alternative (and higher) threshold of at least four participants in community *c* to set $OfferNRPS_{ct}$ being equal to 1. Furthermore, in yet another more stringent definition, we rely on a definition of least seven individuals within the community participating in the NRPS regarding when the community indicator switches from zero (non-participating) to 1 (participating).

We report the results from these additional analyses in Online Appendix Table B4. The results demonstrate that our original estimates are robust to the alternative and more conservative definitions of the threshold, which determines when the variable $OfferNRPS_{ct}$ switches its binary values.

Using Online Administrative Data. Using data from online sources, we perform a final robustness check. For this empirical exercise, we comb data from Chinese newspapers (online or in paper format) based on public announcements regarding geographic participation in the

[64] We conduct an additional extension exercise to address another potential source of measurement error in the variable that measures whether a community implements the NRPS. In the primary analysis, we define the NRPS program's implementation at the community level based on survey data at the individual level. In this empirical approach, if at least one individual in the community reports participating in the NRPS, we define the community as implementing the NRPS. However, communities with very few NRPS participants may be systematically similar in unobservable factors we cannot observe. This scenario could produce a measurement error affecting our measurement of the instrument. Therefore, in an additional extension exercise, we re-estimate by removing communities that report very few NRPS participants within their boundaries. We then proceed by re-estimating the main specifications reported in Section III. Online Appendix B Table B5 reports the results for this extension exercise. The pattern of the results remains consistent with the main results.

NPRS. The two levels for which such data on public announcements are available are at the city and community levels. We specifically focus on the public announcements for NRPS implementation in Heilongjiang Province.[65] Based on these public announcements, we can identify whether a city (or communities within a city) participates in the NRPS in a given year from 2009 to 2013, our primary analysis period. We can also identify the exact timing of when specific cities (and communities within these cities) switched from non-participation to participation in the NRPS.

However, we face a challenge related to how the CHARLS defines the *community* unit and administrative units available from public announcements. We cannot map the actual communities (within cities) to the community units (i.e., the variable *community ID*) in the CHARLS survey, the primary analysis unit in our analyses. This mismatch is due to an inconsistent definition of the "community" variable in the CHARLS and the administrative *community unit* (available in online records). Due to this mismatch, we cannot re-estimate the main specifications (performed at the community level), and therefore, we rely instead on the administrative data for community-level participation. However, we can re-estimate our primary specifications at the city level (i.e., at a higher geographic level than the community level used in the primary analysis). We execute this auxiliary exercise at the city level because: (1) we observe the actual number of communities, based on online public announcements, within a city that implemented the NRPS, and (2) we know the total number of communities, a fixed constant, in a city.

[65] In this additional estimation exercise, we can estimate the original specifications at the city level. Specifically, we can compute the treatment intensity (percent of city participating in the NRPS) based on the following formula: $city\_participation_t$=(# communities in a city that offered the$NRPS_{ct}$) (# total communities in a city)$_{ct}$. This additional robustness check's main advantage is that we can observe the number of communities that implement the NRPS program based on public announcements (the numerator). The denominator of the fraction presented above is the total number of communities, and that number is a fixed constant. A significant disadvantage of this approach is that we can re-estimate the specifications from Section III only at the city level (and only for the Heilongjiang Province. We can obtain data on city or community announcements regarding NRPS implementation). This implies that in this additional analysis, the number of observations is low, limiting the statistical power for statistical inference.

Therefore, we redefine our main treatment for this additional empirical exercise: we change the treatment definition from the binary variable used in our main analyses (at the community level) to a continuous variable that measures treatment intensity (at the city level).[66] Based on this reconstructed definition of treatment, we re-estimate this robustness check at the city level instead of the community level, an estimation level for the main analysis.

Using data for Heilongjiang Province, we re-estimate the main specifications outlined in Section III. However, we use a continuously defined treatment variable. We report the results for this final robustness check at the city level in Online Appendix Table B6. This additional analysis (at the city level) relies on a minimal sample. Despite this statistical power limitation, the results echo the pattern reported in the main analysis. The effect size and the direction of the program effects are consistent with our main estimates based on the CHARLS. Therefore, this additional analysis provides bolsters the results from the main analysis.

VI. Concluding Remarks

In this paper, we investigate the effect of a pension scheme on human capital depreciation in the form of cognitive decline among the elderly in rural China. By using new longitudinal data available from the CHARLS for older individuals, we examine the effects on two categories of cognitive functioning among the elderly: episodic memory and intact mental status. We find large and significant adverse impacts of the pension program on cognition outcomes. The estimated program impacts are similar to other negative findings in the context of high-income countries, such as the US, England, and the European Union (Rohwedder and Willis 2010;

[66] The Heilongjiang province is ideal for this empirical exercise for several reasons. First, online announcements regarding city-level implementations of the NRPS are readily available regarding NRPS implementation between 2011 and 2013 at the city level. As outlined in Section II, data is available from the two CHARLS waves for this period. Second, the province is one of the largest provinces in China. This factor can considerably facilitate the re-estimation exercise because our primary empirical approach relies on identifying variation based on time and space. Third, most of the city-level implementation of the NRPS in this province occurred around 2013. Other cities or areas, within other provinces, had either already adopted the NRPS before 2013, or no information for NRPS implementation at the city-level was available online.

Mazzonna and Peracchi 2012). Individuals in areas that implement the NRPS score considerably lower than individuals who reside in areas that do not offer the NRPS program.

We find substantially larger program impacts on the cognition measure that tests delayed word recall. Previous neurological research documents the importance of this measure, particularly in detecting the difference between normal aging among the elderly and individuals more likely to experience an earlier onset of dementia.

Furthermore, our findings support the *mental retirement* hypothesis, the idea that decreased mental activity atrophies cognitive skills. We demonstrate that retirement policies can play a significant role in explaining the cognitive decline in old age. However, future studies can shed light on how specific activities affect cognition. Two additional areas will be of interest regarding the nexus between retirement and cognitive decline in developing countries. First, what role does the type of job—formal versus informal or white-collar versus blue-collar—play in determining individual mental decline speed? Second, it is essential to uncover and examine the underlying mechanisms between retirement and cognitive decline. A crucial mediating factor in developing countries is the role of informal social networks, social status, and the frequency and quality of social interactions.

Finally, our findings have implications that call for closer examination of the role retirement programs can play in accelerating human capital depreciation in late adulthood. Policy interventions targeting the elderly can have powerful economic consequences. Cognitive impairments among the elderly, even if not severely debilitating, bring about a loss of quality of life and can have negative welfare consequences. Policies aimed at slowing down the cognitive decline in older ages are likely to generate large positive spillovers.

References

Ackerman, Phillip L., Margaret E. Beier, and Mary O. Boyle. 2005. "Working Memory and Intelligence: The Same or Different Constructs?" *Psychological Bulletin* 131 (1): 30–60. https://doi.org/10.1037/0033-2909.131.1.30.

Andrews, Martyn, Thorsten Schank, and Richard Upward. 2006. "Practical Fixed-Effects Estimation Methods for the Three-Way Error-Components Model." *The Stata Journal* 6 (4): 461–81. https://doi.org/10.1177/1536867X0600600402.

Autor, David H. 2003. "Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing." *Journal of Labor Economics* 21 (1): 1–42. https://doi.org/10.1086/344122.

Banks, James, and Zoe Oldfield. 2007. "Understanding Pensions: Cognitive Function, Numerical Ability and Retirement Saving." *Fiscal Studies* 28 (2): 143–70. https://doi.org/10.1111/j.1475-5890.2007.00052.x.

Bingley, Paul, and Alessandro Martinello. 2013. "Mental Retirement and Schooling." *European Economic Review* 63 (October 2013): 292–98. https://doi.org/10.1016/j.euroecorev.2013.01.004.

Bloom, David E., and Roddy McKinnon. 2014. "International Handbook on Ageing and Public Policy: The Design and Implementation of Pension Systems in Developing Countries: Issues and Options." *Institute for the Study of Labor, IZA Policy Paper No. 59*. Bonn, Germany: Institute for the Study of Labor. http://ftp.iza.org/pp59.pdf.

Bonsang, Eric, Stphane Adam, and Sergio Perelman. 2012. "Does Retirement Affect Cognitive Functioning?" *Journal of Health Economics* 31 (3): 490–501. https://doi.org/10.1016/j.jhealeco.2012.03.005.

Brown, Richard E. 2016. "Hebb and Cattell: The Genesis of the Theory of Fluid and Crystallized Intelligence." *Frontiers in Human Neuroscience* 10. https://doi.org/10.3389/fnhum.2016.00606.

Cai, Fang, John Giles, Philip O'Keefe, and Dewen Wang. 2012. *The Elderly and Old Age Support in Rural China*. 1st ed. Book, Whole. Washington, DC: World Bank Publications.

Coe, Norma B., Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen Maurer. 2012. "The Effect of Retirement on Cognitive Functioning." *Health Economics* 21 (8): 913–27. https://doi.org/10.1002/hec.1771.

Cutler, David M., and Richard Johnson. 2004. "The Birth and Growth of the Social Insurance State: Explaining Old Age and Medical Insurance Across Countries." *Public Choice* 120 (1–2): 87–121. https://link.springer.com/content/pdf/10.1023/B:PUCH.0000035859.20258.e0.pdf.

Dorfman, Mark C., Robert Holzmann, Philip O'Keefe, Dewen Wang, Yvonne Sin, and Richard Hinz. 2013. *China's Pension System: A Vision*. 1st ed. Book, Whole. Washington, DC: World Bank Publications.

Eibich, Peter. 2015. "Understanding the Effect of Retirement on Health: Mechanisms and Heterogeneity." *Journal of Health Economics* 43 (September 2015): 1–12.

https://doi.org/10.1016/j.jhealeco.2015.05.001.
Feldstein, Martin, and Jeffrey B. Liebman. 2002. "Chapter 32: Social Security." In *Handbook of Public Economics*, edited by Alan J. Feldstein Auerbach and Martin, 1st ed., 4:2245–2324. Amsterdam, The Netherlands: Elsevier. https://doi.org/10.1016/S1573-4420(02)80011-8.
Fillenbaum, G. G., D. C. Hughes, A. Heyman, L. K. George, and D. G. Blazer. 1988. "Relationship of Health and Demographic Characteristics to Mini-Mental State Examination Score among Community Residents." *Psychological Medicine* 18 (3): 719–26. https://doi.org/10.1017/s0033291700008412.
Grip, Andries de, Arnaud Dupuy, Jelle Jolles, and Martin van Boxtel. 2015. "Retirement and Cognitive Development in the Netherlands: Are the Retired Really Inactive?" *Economics & Human Biology* 19 (December 2015): 157–69. https://doi.org/10.1016/j.ehb.2015.08.004.
Gruber, Jonathan. 1994. "The Incidence of Mandated Maternity Benefits." *The American Economic Review* 84 (3): 622–41. https://www.jstor.org/stable/2118071.
Heckman, James J. 2000. "Policies to Foster Human Capital." *Research in Economics* 54 (1): 3–56. https://doi.org/10.1006/reec.1999.0225.
Hudson, Sally, Peter Hull, and Liebersohn, Jack. 2017. "Interpreting Instrumented Difference-in-Difference." Massachusetts Institute of Technology Mimeo. Cambridge, MA: Massachusetts Institute of Technology. http://www.mit.edu/~hull/DDIV.pdf.
Jones, Lauren, and Tansel Yilmazer. 2018. "The Long-Run Effect of the Earned Income Tax Credit on Cognitive Ability." In *Summer Research Workshop*. Madison, Wisconsin: University of Wisconsin, The Institute for Research on Poverty. https://www.irp.wisc.edu/newsevents/workshops/SRW/2018/participants/presentations/14-Jones-IRP2018-presentations.pdf.
Katz, Lawrence F. 1996. "Wage Subsidies for the Disadvantaged." NBER Working Paper No. 5679. *National Bureau of Economic Research, Mimeo*. Vol. No. w5679. Cambridge, MA: National Bureau of Economic Research, Mimeo. https://www.nber.org/papers/w5679.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119. https://doi.org/10.1111/j.1468-0262.2007.00733.x.
Korniotis, George M., and Alok Kumar. 2011. "Do Older Investors Make Better Investment Decisions?" *The Review of Economics and Statistics* 93 (1): 244–65. https://doi.org/10.1162/REST_a_00053.
Laakso, Mikko P., Merja Hallikainen, Tuomo Hänninen, Kaarina Partanen, and Hilkka Soininen. 2000. "Diagnosis of Alzheimer's Disease: MRI of the Hippocampus vs Delayed Recall." *Neuropsychologia* 38 (5): 579–84. https://doi.org/10.1016/s0028-3932(99)00111-6.
Langa, Kenneth M., Lindsay H. Ryan, Ryan J. McCammon, Richard N. Jones, Jennifer J. Manly, Deborah A. Levine, Amanda Sonnega, Madeline Farron, and David R. Weir. 2020. "The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study Design and Methods." *Neuroepidemiology* 54 (1): 64–74. https://doi.org/10.1159/000503004.
Lei, Xiaoyan, Chuanchuan Zhang, and Yaohui Zhao. 2013. "Incentive Problems in China's New Rural Pension Program." *Research in Labor Economics* 37 (Journal Article): 181–201.

https://doi.org/10.1108/S0147-9121(2013)0000037010.
Liu, Tao, and Li Sun. 2016. “Pension Reform in China.” *Journal of Aging & Social Policy* 28 (1): 15–28. https://doi.org/10.1080/08959420.2016.1111725.
Mazzonna, Fabrizio, and Franco Peracchi. 2012. “Ageing, Cognitive Abilities and Retirement.” *European Economic Review* 56 (4): 691–710. https://doi.org/10.1016/j.euroecorev.2012.03.004.
McArdle, John J, James P Smith, and Robert Willis. 2011. “Cognition and Economic Outcomes in the Health and Retirement Survey.” In *Explorations in the Economics of Aging*, edited by David Wise, 209–36. Chicago, IL: University of Chicago Press.
Nikolov, Plamen, and Alan Adelman. 2019a. “Do Private Household Transfers to the Elderly Respond to Public Pension Benefits?” *The Journal of the Economics of Ageing* 14 (1): Article 100204. https://doi.org/10.1016/j.jeoa.2019.100204.
———. 2019b. “Short-Run Health Consequences of Retirement and Pension Benefits: Evidence from China.” *Forum for Health Economics and Policy* 21 (2): 1–27. https://doi.org/10.1515/fhep-2017-0031.
Ofstedal, Mary Beth, Gwenith G. Fisher, and A. Regula Herzog. 2005. “Documentation of Cognitive Functioning Measures in the Health and Retirement Study.” *University of Michigan, HRS Documentation Report DR-006*. Vol. 10. Ann Arbor, MI: University of Michigan. http://hrsonline.isr.umich.edu/sitedocs/userg/dr-006.pdf.
Olden, Andreas, and Jarle Moen. 2020. “The Triple Difference Estimator.” 1550–4066. Institutt for Foretaksøkonomi, Discussion Papers. Bergen, Norway: NHH Norwegian School of Economics. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3582447.
Prull, Matthew W., John DE Gabrieli, and Silvia A. Bunge. 2000. “Age-Related Changes in Memory: A Cognitive Neuroscience Perspective.” In *Handbook of Aging and Cognition*, edited by Fergus. Salthouse Craik Timothy A., 2nd ed., 91–153. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
Quan, Lu. 2012. “Analysing the Coverage Gap in China.” *London, HelpAge International*. London, UK: HelpAge International. http://www.pension-watch.net/knowledge-centre/?guid=51d5332022948&order=n.
Rohwedder, Susann, and Robert J. Willis. 2010. “Mental Retirement.” *The Journal of Economic Perspectives* 24 (1): 119–38. https://doi.org/10.1257/jep.24.1.119.
Sabates-Wheeler, Rachel, and Johannes Koettl. 2010. “Social Protection for Migrants: The Challenges of Delivery in the Context of Changing Migration Flows.” *International Social Security Review* 63 (3-4): 115–44. https://doi.org/10.1111/j.1468-246X.2010.01372.x.
Society of Actuaries. 2016. “The Future of Retirement in China: Retirement Definitions and Demographics.” 009807–0516. Schaumburg, IL: SOA. https://www.soa.org/globalassets/assets/Files/Research/Projects/research-2016-retirement-definitions-demographics.pdf.
Souchay, Cline, Michel Isingrini, and Laurence Espagnet. 2000. “Aging, Episodic Memory Feeling-of-Knowing, and Frontal Functioning.” *Neuropsychology* 14 (2): 299. https://doi.org/10.1037//0894-4105.14.2.299.
The World Bank. 2017. “Population Growth.” Population Growth Report. *The World Bank, Population Growth Report*. United Nations Population Division. World Population

Prospects: 2019 Revision. Washington, DC: The World Bank. https://data.worldbank.org/indicator/SP.POP.GROW?page=6.
UNC Chapel Hill, Carolina Population Center. 2015. “Map of Survey Regions.” Map of Survey Regions. *University of North Carolina*. Chapel Hill, NC: University of North Carolina. https://www.cpc.unc.edu/projects/china/about/proj_desc/chinamap.
UNC Chapel Hill, Carolina Population Center, CHNS Research Team. 2010. “Survey Design.” *University of North Carolina*. Chapel Hill, NC: University of North Carolina. https://www.cpc.unc.edu/projects/china/about/design/survey.
United Nations. 2019. “World Population Prospects 2019: Highlights.” *United Nations, Department of Economic and Social Affairs*. United Nations, Department of Economic and Social Affairs, Population Division. New York, NY: United Nations. https://population.un.org/wpp/Publications/Files/WPP2019_10KeyFindings.pdf.
Vilela, Andrea. 2013. “Pension Coverage in China and the Expansion of the New Rural Social Pension.” Briefing no. 11. *HelpAge International*. Vol. ISBN 1872590799. London, UK: HelpAge International. https://www.refworld.org/pdfid/5301df5d4.pdf.
Welsh, Kathleen, Nelson Butters, James Hughes, Richard Mohs, and Albert Heyman. 1991. “Detection of Abnormal Memory Decline in Mild Cases of Alzheimer’s Disease Using CERAD Neuropsychological Measures.” *Archives of Neurology* 48 (3): 278–81. https://doi.org/10.1001/archneur.1991.00530150046016.

Figures and Tables

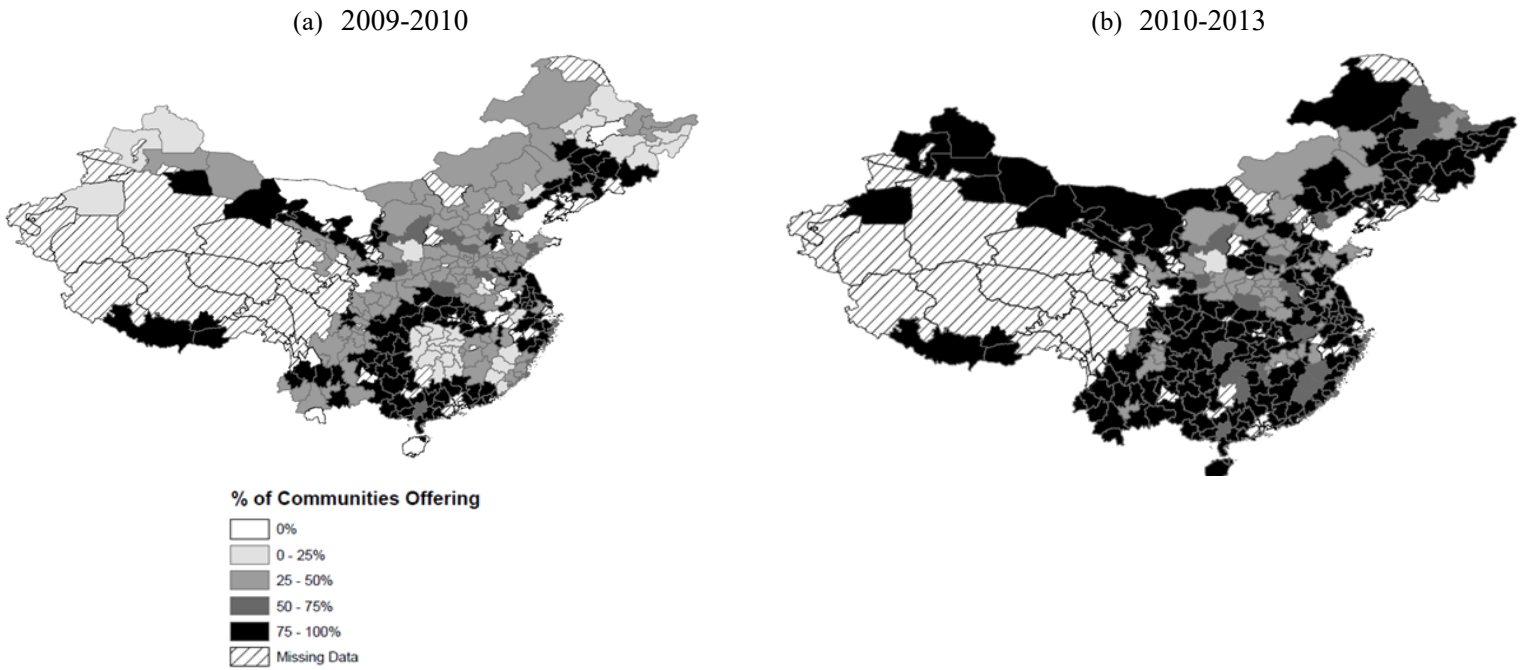


Fig 1. Geographic Implementation of NRPS. This figure shows the timely implementation of NRPS. "% of Communities Offering" indicates the percent of communities (*shequs*) within the province that implemented the NRPS.

Table 1: Summary Statistics.

		Baseline		
	Full Sample	NRPS Participants	NRPS Non-Participants	p-value[a]
Demographics of Respondents				
Respondent's Age	59.31 (10.01)	58.43 (9.68)	58.44 (10.24)	0.99
# of Household Residents	3.74 (1.87)	3.68 (1.78)	3.75 (1.88)	0.04
# Living Children	2.77 (1.44)	2.81 (1.39)	2.74 (1.45)	0.07
Percent Female	0.53 (0.50)	0.54 (0.50)	0.53 (0.50)	0.38
Percent Married	0.80 (0.40)	0.81 (0.39)	0.78 (0.41)	0.00
Percent Living Near Children	0.90 (0.30)	0.91 (0.28)	0.92 (0.27)	0.40
Percent With At Least Lower Secondary Education	0.48 (0.50)	0.48 (0.50)	0.46 (0.50)	0.10
Labour Market and Health Outcomes				
Weekly Work Hours	45.45 (23.87)	47.26 (24.07)	46.89 (22.70)	0.50
Percent Currently Working	0.70 (0.46)	0.70 (0.46)	0.69 (0.46)	0.11
Percent Working in Agriculture	0.72 (0.45)	0.72 (0.45)	0.73 (0.45)	0.49
Percent Reporting Poor/Fair Health	0.25 (0.43)	0.27 (0.44)	0.26 (0.44)	0.23
Respondent's BMI	23.40 (3.84)	23.62 (3.91)	23.05 (3.81)	0.00
Percent Visited Doctor (Past Month)	0.20 (0.40)	0.20 (0.40)	0.19 (0.39)	0.08
Percent Stayed in Hospital (Past Year)	0.11 (0.31)	0.10 (0.29)	0.09 (0.28)	0.06
Percent Ever Smoked	0.41 (0.49)	0.40 (0.49)	0.40 (0.49)	0.98
Percent Smoking Now	0.25 (0.44)	0.29 (0.45)	0.30 (0.46)	0.40
Cognition[b]				
Immediate Recall Score	3.79 (1.76)	3.93 (1.69)	3.77 (1.70)	0.00
Delayed Recall Score	2.86 (2.00)	2.91 (1.91)	2.89 (1.96)	0.61
Total Recall Score	6.67 (3.47)	6.85 (3.32)	6.68 (3.36)	0.02
Cognitive Memory Index	0.00 (1.43)	0.06 (1.38)	0.00 (1.39)	0.06
Observations	28,034[c]	10,011	3,680	

Notes: Standard deviations are reported in parenthesis. The full sample consists of observations from 2011 and 2013 waves, whereas baseline observations are for the sub-sample of participants and non-participants from the 2011 wave only. (a) We test the null hypothesis that the difference in participant and non-participant means is equal to 0. (b) Low (or Negative) values denote lower performance on the cognition test. (c) Includes observations from all waves.

Table 2: NRPS Participation and Cognitive Performance.

	Immediate Word Recall[a]	Immediate Word Recall[a]	Delay Word Recall[a]	Delay Word Recall[a]	Total Recall[a]	Total Recall[a]	Cognitive Memory Index[b]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A (ITT):								
Offered NRPS × Above60 [c]	-0.144***	0.134**	-0.230***	-0.272***	-0.353***	-0.407***	-0.103**	-0.123***
	(0.052)	(0.056)	(0.052)	(0.063)	(0.093)	(0.107)	(0.040)	(0.041)
Baseline Mean	3.792	3.792	2.862	2.862	6.678	6.678	0.000	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.230	0.320	0.215	0.318	0.247	0.375	0.313	0.483
Observations	21,202	15,540	21,202	15,540	21,202	15,540	21,202	15,540
Panel B (TOT):								
NRPS Participation × Above60 [d]	-0.208*	-0.270**	-0.425***	-0.547***	-0.633***	-0.816***	-0.212**	-0.247***
	(0.120)	(0.113)	(0.122)	(0.129)	(0.214)	(0.217)	(0.087)	(0.084)
Baseline Mean	3.792	3.792	2.862	2.862	6.678	6.678	0.000	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes
F-Stat (First Stage)	241.242	437.06	241.242	437.06	241.242	437.06	241.242	437.06
R-squared	0.065	0.003	0.152	0.004	0.102	0.005	0.110	0.004
Observations	21,202	15,540	21,202	15,540	21,202	15,540	21,202	15,540

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) Our DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. (d) Individual participation is instrumented with the NRPS availability in the local municipality. Individual level controls: Above60 (1= Yes), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Gender (=1 if Female) and Marital Status (=1 if Married) are included in the specifications without FEs. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community×Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table 3: Retirement and Cognitive Performance.

	Immediate Word Recall[b]	Immediate Word Recall[b]	Delay Word Recall[b]	Delay Word Recall[b]	Total Recall[b]	Total Recall[b]	Cognitive Memory Index[c]	Cognitive Memory Index[c]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Retired (Yes=1)[a]	-0.469*	-0.810***	-0.575	0.703	-0.906	-0.107	-0.529**	-0.556
	(0.283)	(0.189)	(0.369)	(1.478)	(0.613)	(1.570)	(0.252)	(1.220)
Baseline Mean	3.792	3.792	2.862	2.862	6.678	6.678	0.000	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.081	0.010	0.076	0.002	0.100	0.009	0.166	0.004
F-Stat (First Stage)	47.48	23.97	48.18	23.97	47.39	23.97	45.82	23.97
Observations	22,444	15,540	22,329	15,540	22,226	15,540	21,258	15,540

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Directly asked about retirement procedure. "Have you completed retirement procedures (including early retirement) or internal retirement (Retirement from government departments, enterprises and institutions, not including retirement in the sense of getting agricultural insurance)?" A positive answer is coded as being retired. (b) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (c) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. Individual level controls: Above60 (1= Yes), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Gender (=1 if Female) and Marital Status (=1 if Married) are only included in the specifications without FEs. Columns (1) through (4) are estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis. The number of observations in this table differ from Table 2 because of different independent variables used.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table 4: Duration of NRPS Benefits and Cognitive Decline.

	Immediate Word Recall[b]	Immediate Word Recall[b]	Delay Word Recall[b]	Delay Word Recall[b]	Total Recall[b]	Total Recall[b]	Cognitive Memory Index[c]	Cognitive Memory Index[c]
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Offer NRPS × Above 60	-0.047	-1.215	0.375	1.383	0.328	0.168	-0.052	-0.728
	(0.487)	(1.885)	(0.482)	(1.891)	(0.874)	(3.632)	(0.378)	(1.243)
Duration of NRPS Benefits	-0.154	-0.584*	-0.122	-0.629	-0.276	-1.213*	-0.088	-0.647**
(Receiving for 1-2 years)	(0.144)	(0.308)	(0.158)	(0.425)	(0.264)	(0.613)	(0.109)	(0.290)
Duration of NRPS Benefits	0.204	-0.619	0.503	-0.186	0.707	-0.805	0.333	-0.603
(Receiving for 3 or more years)	(0.316)	(0.602)	(0.364)	(0.886)	(0.534)	(1.350)	(0.225)	(0.548)
Offer NRPS (Yes=1) × Above 60 (Yes=1) × Duration of NRPS Benefits (1-2 years)[a]	0.168	1.490	-0.046	0.212	0.122	1.703	0.111	1.292*
	(0.186)	(0.919)	(0.210)	(0.928)	(0.344)	(1.531)	(0.143)	(0.664)
Offer NRPS (Yes=1) × Above 60 (Yes=1) × Duration of NRPS Benefits (3 years or more)[a]	-0.325	1.674	-0.727*	0.070	-1.052*	1.744	-0.382	1.445*
	(0.339)	(1.014)	(0.393)	(1.206)	(0.582)	(1.894)	(0.246)	(0.808)
Above 60	-0.280	-0.398	-0.273	-1.135	-0.554	-1.533	-0.193	-0.393
	(0.419)	(0.998)	(0.382)	(1.326)	(0.725)	(2.229)	(0.322)	(1.092)
Baseline Mean	3.792	3.792	2.862	2.862	6.678	6.678	0.000	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Individual FE	No	Yes	No	Yes	No	Yes	No	Yes
R-squared	0.234	0.781	0.225	0.760	0.252	0.786	0.310	0.818
Observations	22,199	15,540	22,092	15,540	21,992	15,540	21,041	15,540

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Directly asked about retirement procedure. "Have you completed retirement procedures (including early retirement) or internal retirement (Retirement from government departments, enterprises and institutions, not including retirement in the sense of getting agricultural insurance)?" A positive answer is coded as being retired. (b) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (c) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. Individual level controls: Above60 (1= Yes), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Gender (=1 if Female) and Marital Status (=1 if Married) are only included in the specifications without FEs. Columns (1) through (4) are estimated using Ordinary Least Squares (OLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis. The number of observations in this table differ from Table 2 because of different independent variables used.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table 5: Mechanisms Analysis (Labor, Mental and Social Engagements)

	Labor Activities				Mental Stimulation			Social Engagement				
	#Months Worked (Past year)	Hours Daily Worked (Per Week)	Self-Employment #Months Worked (Past year)	Self-Employment Hours Daily Worked (Per Week)	Played Majong Last Month (Yes=1)	Adult Education Course Last Month (Yes=1)	Mental Stimulation Index	Helped Friends Last Month (Yes=1)	Any Community Activity Last Month (Yes=1)	Volunteered Last Month (Yes=1)	Interact w Friends Last Month (Yes=1)	Social Engagement Index
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Panel A (ITT):												
Offered NRPS × Above60 [a]	-0.160	-0.024	-1.342**	-1.224**	-0.011	-0.002	-0.043	-0.023**	-0.026	-0.040***	-0.034***	-0.105**
	(0.464)	(0.034)	(0.565)	(0.516)	(0.010)	(0.001)	(0.028)	(0.009)	(0.016)	(0.012)	(0.012)	(0.040)
Baseline Mean	7.942	3.581	8.876	7.874	0.607	0.032	0.548	0.099	0.095	0.061	0.546	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.381	0.154	0.379	0.373	0.441	0.264	0.350	0.146	0.316	0.180	0.356	0.115
Observations	2,668	15,203	1,748	1,701	12,842	11,134	11,818	21,198	21,198	21,198	21,198	21,198
Panel B (TOT):												
NRPS Participation × Above60 [b]	-0.377	-0.055	-3.151**	-2.968**	-0.025	-0.003	-0.099	-0.052**	-0.050	-0.063***	-0.075***	-0.239**
	(1.095)	(0.079)	(1.346)	(1.263)	(0.022)	(0.003)	(0.065)	(0.021)	(0.033)	(0.019)	(0.028)	(0.094)
Baseline Mean	7.942	3.581	8.876	7.874	0.607	0.032	0.548	0.099	0.095	0.061	0.546	0.000
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-Stat (First Stage)	38.99	195.89	36.97	32.48	238.71	234.67	236.72	291.12	235.37	234.65	239.65	291.12
R-squared	0.380	0.154	0.357	0.357	0.034	0.012	0.030	0.012	0.004	0.006	0.030	0.001
Observations	2,668	15,203	1,748	1,701	12,842	11,134	11,818	21,198	11,178	11,145	14,775	21,198

Notes: The number of observations reflect the number of non-missing values for each outcome in the CHARLS. The table reports estimates of the DDD estimator for the NRPS treatment effect. Food expenses are in constant 2011 Yuan. (a) The DDD estimator (NRPS availability interacted with an indicator for being over 60 years old). The control group is individuals under the age of 60 living in eligible communities that didn't offer NRPS between 2011 and 2013. (b) Individual participation instrumented with the policy variable. Individual level controls: Above60 (1= Yes), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Gender (=1 if Female) and Marital Status (=1 if Married) are only included in the specifications without FEs. Regular alcohol drinker: drinking at least once per week in the last year. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis. In constant 2011 Yuan. *p< 0.10, **p< 0.05, ***p< 0.01.

Table 6: Mechanisms Analysis (Health Behaviors).

	Health Behaviors and Nutrition				
	Hrs Sleep per Night (Last Year)	Currently Smoking (Yes=1)	Regular Alcohol Drinker (Yes=1)	Health Behaviors Index	HH Food Expenses (Last week in Yuan)
	(1)	(2)	(3)	(4)	(5)
Panel A (ITT):					
Offered NRPS × Above60 [a]	0.157**	-0.019*	-0.020**	0.083**	17.434
	(0.061)	(0.011)	(0.009)	(0.035)	(39.528)
Baseline Mean	6.281	0.254	0.186	0.000	192.443
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.107	0.389	0.257	0.114	0.052
Observations	20,913	18,965	20,688	18,199	23,822
Panel B (TOT):					
NRPS Participation × Above60 [b]	0.354**	-0.042*	-0.045**	0.185**	37.767
	(0.142)	(0.025)	(0.020)	(0.079)	(85.542)
Baseline Mean	6.281	0.254	0.186	0.000	178.488
Controls	Yes	Yes	Yes	Yes	Yes
F-Stat (First Stage)	244.47	223.79	243.02	296.37	244.743
R-squared	0.104	0.389	0.256	0.012	0.053
Observations	20,913	18,965	20,688	18,199	23,822

Notes: The number of observations reflect the number of non-missing values for each outcome in the CHARLS. The table reports estimates of the DDD estimator for the NRPS treatment effect. Food expenses are in constant 2011 Yuan. (a) The DDD estimator (NRPS availability interacted with an indicator for being over 60 years old). The control group is individuals under the age of 60 living in eligible communities that didn't offer NRPS between 2011 and 2013. (b) Individual participation instrumented with the policy variable. Individual level controls: Above60 (1= Yes), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Gender (=1 if Female) and Marital Status (=1 if Married) are included in the specifications without FEs. Regular alcohol drinker: drinking at least once per week in the last year. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis. In constant 2011 Yuan. *p< 0.10, **p< 0.05, ***p< 0.01.

Table 7: Heterogeneous Treatment Effects by Gender.

	Immediate Word Recall[a]	Delayed Word Recall[a]	Total Recall[a]	Cognitive Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
Offered NRPS × Above60 × Female [c]	-0.092	0.073	-0.019	-0.008
	(0.102)	(0.125)	(0.195)	(0.075)
Offered NRPS × Above60	-0.036	-0.249**	-0.285*	-0.070
	(0.076)	(0.086)	(0.141)	(0.055)
Offered NRPS × Female	-0.005	-0.100	-0.105	-0.025
	(0.065)	(0.085)	(0.127)	(0.047)
Above60 × Female	0.037	-0.082	-0.045	-0.104*
	(0.079)	(0.092)	(0.150)	(0.061)
Above60	-0.429***	-0.407***	-0.836***	-0.268***
	(0.058)	(0.066)	(0.110)	(0.045)
Female	0.126**	0.276***	0.402***	0.025
	(0.050)	(0.059)	(0.091)	(0.035)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.223	0.217	0.244	0.304
Observations	21,202	21,202	21,202	21,202
Panel B (TOT):				
NRPS Participation × Above60 × Female[d]	-0.209	0.152	-0.057	-0.021
	(0.228)	(0.280)	(0.436)	(0.167)
NRPS Participation × Above60[d]	-0.076	-0.452***	-0.528**	-0.147
	(0.168)	(0.176)	(0.305)	(0.118)
NRPS Participation × Female	-0.004	-0.097	-0.101	-0.024
	(0.065)	(0.085)	(0.127)	(0.047)
Above60 × Female	0.081	-0.101	-0.020	-0.097
	(0.112)	(0.135)	(0.212)	(0.085)
Above60	-0.415***	-0.308***	-0.723***	-0.240***
	(0.081)	(0.094)	(0.153)	(0.061)
Female	0.124**	0.273***	0.397***	0.024
	(0.050)	(0.059)	(0.091)	(0.035)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.110	0.100	0.124	0.193
F-Stat (First Stage)	141.13	141.13	141.13	141.13
Observations	21,202	21,202	21,202	21,202

Notes: The table reports estimates of the DDDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. (d) Individual participation is instrumented with the NRPS availability in the local municipality. Individual level controls: Age, Age Squared, Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. The ITT effects are estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Appendix A

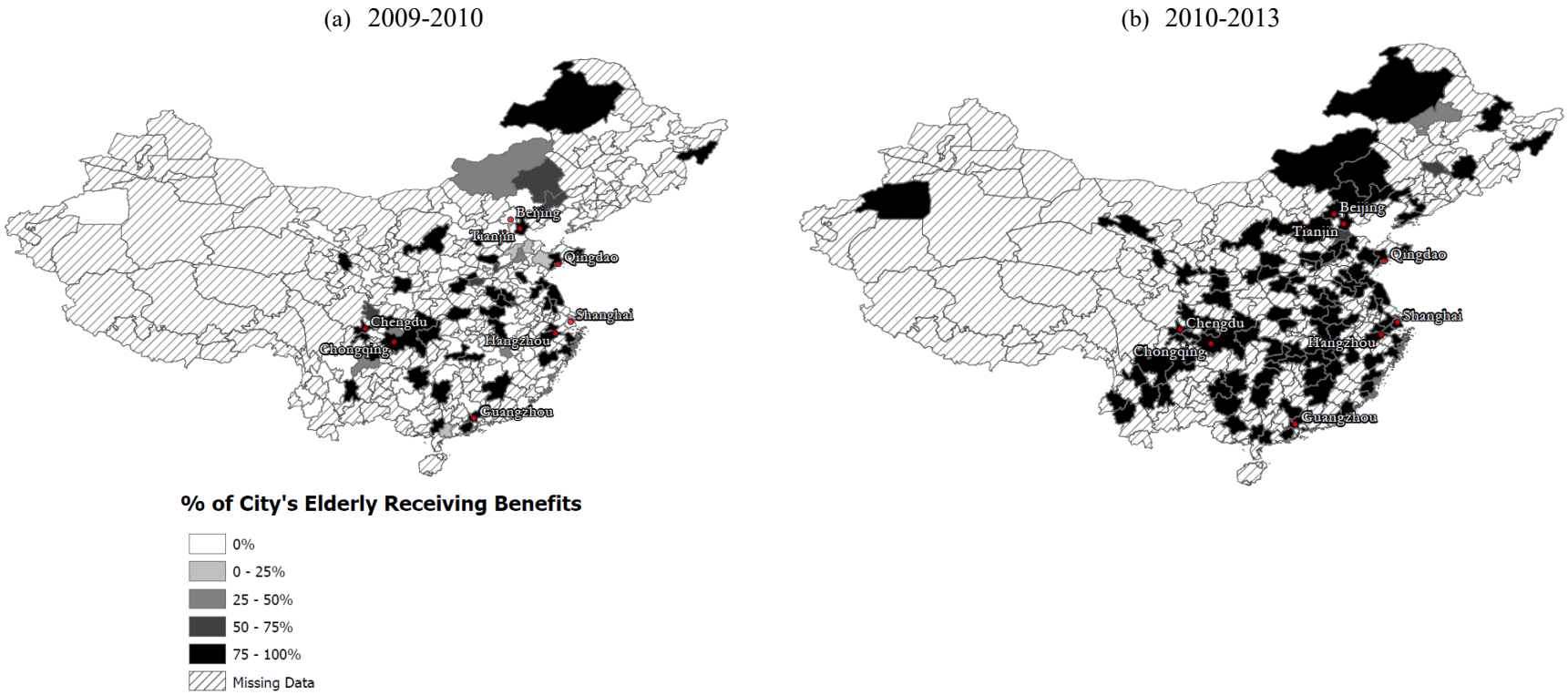


Fig A1. Geographic Implementation of NRPS. This figure reports the percent of the elderly receiving NRPS benefits as a fraction of all elderly in each community. "% of City's Elderly Receiving Benefits Offering" = the number of elderly receiving NRPS benefits in a community (*shequ*)/the number of elderly in a community.

Table A1: Test of Common Trends Using CHNS Data.

		Immediate Word Recall[a]	Delayed Word Recall[a]	Total Recall[a]
		(1)	(2)	(3)
50% Coverage Rate Threshold	Treatment × Age>Above 60 (Yes=1) × 2000	-0.216	-0.286	-0.808
		(0.387)	(0.399)	(0.761)
	Treatment × Age>Above 60 (Yes=1) × 2006	-0.446	-0.325	-0.708
		(0.306)	(0.383)	(0.586)
	R-Squared Adj	0.215	0.228	0.241
	Year FE	Yes	Yes	Yes
	Community FE	Yes	Yes	Yes
	Observations	4,742	4,719	4,615
70% Coverage Rate Threshold	Treatment × Age>Above 60 (Yes=1) × 2000	-0.186	0.519	0.364
		(0.417)	(0.465)	(0.837)
	Treatment × Age>Above 60 (Yes=1) × 2006	-0.479	0.173	-0.281
		(0.309)	(0.383)	(0.660)
	R-Squared Adj	0.214	0.229	0.241
	Year FE	Yes	Yes	Yes
	Community FE	Yes	Yes	Yes
	Observations	4,742	4,719	4,615

Notes: Source: CHNS 2000, 2004, and 2006 Waves. Base year is 2004. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table A2: Mechanisms Analysis By Gender.

	Labor Activities				Mental Stimulation		Social Engagement				Health Behaviors And Nutrition			
	#Months Worked (Past year)	Hrs Daily Worked (Per Week)	Self-Empl #Months Worked (Past year)	Self-Empl Hrs Daily Worked (Week)	Played Majong Last Month (Yes=1)	Adult Education Course Last Month (Yes=1)	Helped Friends Last Month (Yes=1)	Any Community Activity Last Month (Yes=1)	Volunteered Last Month (Yes=1)	Interact w Friends Last Month (Yes=1)	Hrs Sleep Night (Last Year)	Currently Smoking (Yes=1)	Regular Alcohol Drinker (Yes=1)	HH Food Expenses (Last week in Yuan)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A (ITT):														
Offered NRPS ×	0.049	0.006	-1.502	-1.235	0.069**	-4.695*	-0.153	-0.004	0.014	0.064***	0.032	0.006	0.003	7.378
Above60 × Female [c]	(1.007)	(0.065)	(0.935)	(0.951)	(0.032)	(2.572)	(0.231)	(0.022)	(0.022)	(0.024)	(0.068)	(0.026)	(0.024)	(8.714)
Offered NRPS ×	-0.162	-0.032	-0.703	-0.704	-0.165***	2.084*	-0.090	-0.021	-0.047**	-0.066***	0.134	-0.031	-0.018	12.785
Above60	(0.548)	(0.041)	(0.630)	(0.641)	(0.027)	(1.193)	(0.109)	(0.021)	(0.018)	(0.019)	(0.086)	(0.023)	(0.019)	(38.120)
Offered NRPS ×	0.320	-0.068*	-0.750	-0.392	-0.246***	4.407*	0.123	-0.033**	-0.039**	0.010	-0.118*	0.096***	0.018	-32.438
Female	(0.506)	(0.035)	(0.369)	(0.415)	(0.026)	(2.322)	(0.222)	(0.017)	(0.016)	(0.018)	(0.066)	(0.018)	(0.017)	(24.962)
Above60 × Female	-1.382*	-0.042	0.534	0.417	-0.017	4.727*	0.153	0.005	-0.004	-0.006	-0.263***	0.120***	0.124***	-5.829
	(0.741)	(0.044)	(0.649)	(0.650)	(0.016)	(2.572)	(0.228)	(0.011)	(0.016)	(0.004)	(0.083)	(0.020)	(0.020)	(11.579)
Above60	-0.289	-0.224***	-0.250	-1.015*	0.019*	-2.104*	-0.079	0.015	0.019	-0.004	-0.142**	-0.082***	-0.109***	-62.074
	(0.427)	(0.029)	(0.466)	(0.573)	(0.010)	(1.193)	(0.105)	(0.012)	(0.013)	(0.004)	(0.065)	(0.017)	(0.016)	(39.542)
Female	-0.668	-0.122***	0.011	0.213	0.052***	-4.435*	-0.163	0.005	0.003	0.019***	-0.068	-0.594***	-0.483***	0.831
	(0.426)	(0.025)	(0.343)	(0.320)	(0.009)	(2.322)	(0.222)	(0.004)	(0.009)	(0.005)	(0.062)	(0.016)	(0.015)	(5.158)
Baseline Mean	7.942	3.581	8.876	7.874	0.607	0.032	0.548	0.095	0.061	0.546	6.281	0.254	0.33	192.443
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.381	0.155	0.382	0.375	0.441	0.266	0.350	0.316	0.180	0.357	0.108	0.393	0.290	0.053
Observations	2,668	15,203	1,748	1,701	12,842	11,134	11,818	11,178	11,145	14,775	20,913	18,965	21,192	23,835
Panel B (TOT):														
Offered NRPS ×	0.075	0.013	-3.645	-3.393	0.162*	-7.301*	-0.521	-0.017	0.023	0.142***	0.075	0.014	0.007	16.584
Above60 × Female [c]	(2.412)	(0.145)	(2.744)	(2.661)	(0.091)	(3.998)	(0.734)	(0.038)	(0.034)	(0.054)	(0.250)	(0.057)	(0.055)	(18.924)
Offered NRPS ×	-0.372	-0.073	-1.726	-1.698	-0.419***	3.261*	-0.247	-0.039	-0.075**	-0.145***	0.301	-0.071	-0.041	29.194
Above60	(1.283)	(0.095)	(1.485)	(1.501)	(0.080)	(1.863)	(0.329)	(0.035)	(0.029)	(0.043)	(0.196)	(0.052)	(0.043)	(83.065)
Offered NRPS ×	0.325	-0.067*	-0.760	-0.387	-0.245***	4.399*	0.126	-0.032*	-0.039**	0.011	-0.119*	0.096***	0.018	-5.545
Female	(0.508)	(0.035)	(0.464)	(0.413)	(0.027)	(2.317)	(0.219)	(0.017)	(0.016)	(0.018)	(0.066)	(0.018)	(0.017)	(5.717)
Above60 × Female	-1.407	-0.045	1.343	1.179	-0.046	4.672*	0.338	0.012	-0.004	-0.030**	-0.283**	0.119***	0.123***	3.733
	(1.036)	(0.066)	(1.171)	(1.159)	(0.038)	(2.538)	(0.462)	(0.011)	(0.016)	(0.012)	(0.120)	(0.026)	(0.027)	(7.890)
Above60	-0.226	-0.210***	0.093	-0.670	0.117***	-2.076*	-0.015	0.018	0.018	0.029***	-0.198**	-0.069***	-0.101***	-67.125
	(0.616)	(0.044)	(0.661)	(0.796)	(0.031)	(1.174)	(0.202)	(0.013)	(0.013)	(0.010)	(0.095)	(0.023)	(0.021)	(53.143)
Female	-0.666	-0.122***	-0.010	0.193	0.052***	-4.433*	-0.166	0.004	0.003	0.019***	-0.067	-0.594***	-0.483***	0.869
	(0.426)	(0.025)	(0.344)	(0.320)	(0.009)	(2.321)	(0.220)	(0.004)	(0.009)	(0.005)	(0.062)	(0.016)	(0.015)	(5.098)
Baseline Mean	7.942	3.581	8.876	7.874	0.607	0.032	0.548	0.095	0.061	0.546	6.281	0.254	0.33	192.443
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
F-statistic	9.273	110.355	11.792	9.269	156.032	152.785	155.430	153.873	152.811	158.993	156.983	157.600	158.416	169.078
R-squared	0.382	0.155	0.353	0.348	0.433	0.266	0.332	0.316	0.178	0.354	0.106	0.395	0.292	0.053
Observations	2,668	15,203	1,748	1,701	12,842	11,134	11,818	11,178	11,145	14,775	20,913	18,965	21,192	23,835

Notes: The table reports estimates of the DDDD estimator for the NRPS treatment effect. Food expenses are in constant 2011 Yuan. (a) The DDD estimator (NRPS availability interacted with an indicator for being over 60 years old). The control group is individuals under the age of 60 living in eligible communities that didn't offer NRPS between 2011 and 2013. (b) Individual participation instrumented with the policy variable. Individual level controls: Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Regular alcohol drinker: drinking at least once per week in the last year. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis. ***Significant at the 1 percent level. **Significant at the 5 percent level. *Significant at the 10 percent level.

Table A3: Falsification Test Using Placebo Sample.

	Immediate Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)
Offered NRPS × Above60[c]	-0.098	-0.710	-0.183
	(0.245)	(0.494)	(0.207)
Baseline Mean	0.253	0.000	0.000
Controls	Yes	Yes	Yes
R-squared	0.611	0.625	0.620
Observations	604	594	576

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) Our DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. A significant coefficient suggests the differential treatment towards urban pensioners in treated communities relative to urban pensioner in control communities; a cause of concern for the instrument's validity. Individual level controls: Above60 (1=Yes), Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. The specifications are estimated with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table A4: Test on Placebo Outcomes for Specifications (1) and (3).

	Han (=1 if yes)	# Dead Daughter	Mother's Education	# of Living Sons
	(1)	(2)	(3)	(4)
Panel A (ITT):				
Offered NRPS × Above60 [a]	-0.004	-0.013	-0.015	0.010
	(0.004)	(0.029)	(0.012)	(0.024)
Baseline Mean	0.920	1.299	1.190	1.466
Controls	Yes	Yes	Yes	
R-squared	0.652	0.165	0.130	0.235
Observations	20,102	21,202	19,656	21,202
Panel B (TOT):				
NRPS Participation × Above60 [b]	-0.010	-0.032	-0.035	0.025
	(0.009)	(0.071)	(0.030)	(0.059)
Baseline Mean	0.920	1.299	1.190	1.466
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	282.617	279.6213	291.9617	279.6213
R-squared	0.652	0.165	0.130	0.235
Observations	20,102	21,202	19,656	21,202

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) DDD coefficient(NRPS availability interacted with an indicator for being over 60 years old). The control group becomes individuals under the age of 60 living in eligible communities that didn't offer NRPS between 2011 and 2013. (b) Individual participation instrumented with the policy variable. Individual level controls: Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

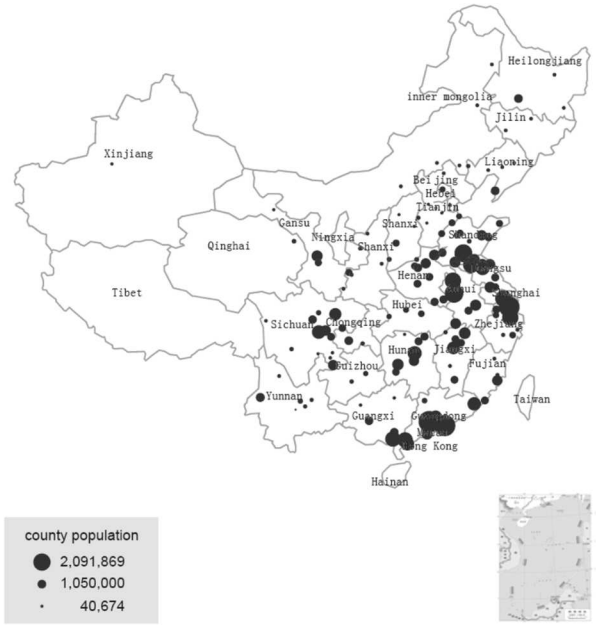
***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Supplementary Tables
Online Appendix B

China Health and Retirement Longitudinal Study



China Health and Nutrition Survey

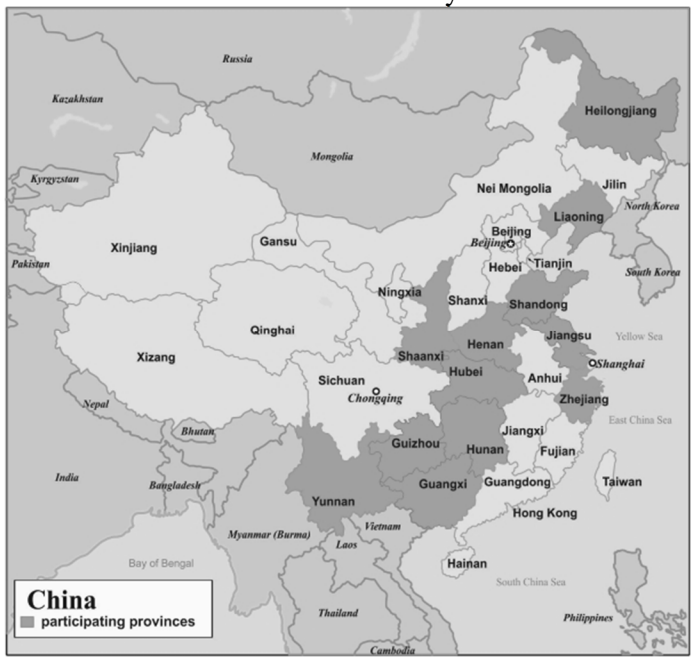


Fig B1. Coverage Maps. Source: China Center for Economic Research (2013) and UNC-Carolina Population Center (2015).

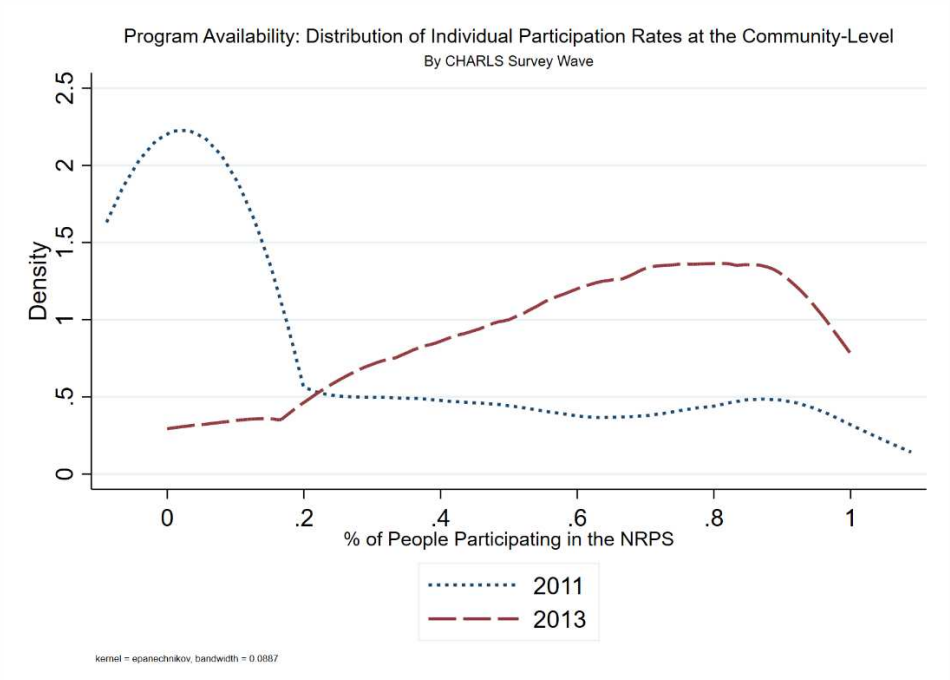


Fig B2. Program Availability: Distribution of Individual Participation Rates at the Community-Level. Source: CHARLS Survey.

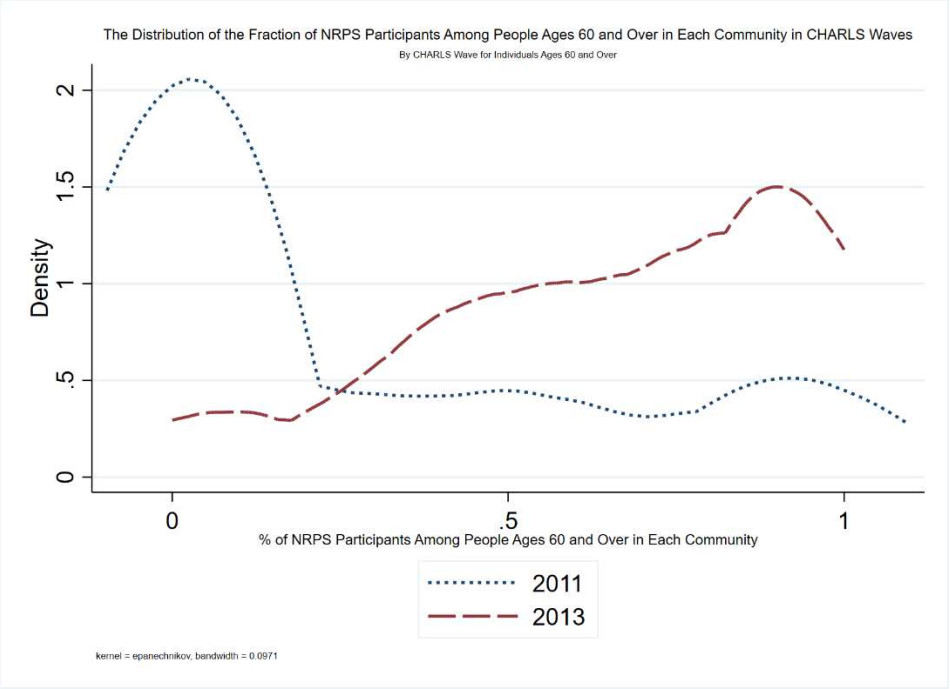


Fig B3. Program Availability: The distribution of the fraction of over 60 people in each community who indicate NRPS participation in each CHARLS wave. Source: CHARLS Survey.

Table B1: Demographic Characteristics between CHARLS and CHNS Datasets (Age 45 and above and Living in Rural Districts).

Variables	CHARLS			CHNS		
	# Observations	Mean	Std. Dev.	# Observations	Mean	Std. Dev.
	(1)	(2)	(3)	(4)	(5)	(6)
Age	30,590	59.27	10.01	17,130	57.65	11.61
Han Nationality	27,280	0.922	0.268	17,086	0.847	0.360
Female	30,644	0.522	0.500	17,136	0.523	0.500
Currently Married	27,309	0.796	0.403	15,870	0.842	0.364
At Least Lower Secondary Education	30,584	0.473	0.499	15,921	0.371	0.483
# of Children	27,346	2.784	1.45	15,080	2.170	1.059
Household Income (Yuan)	19,701	25800.67	64097.47	15,913	27440.65	39603.94

Notes: Yuan reported are in in constant 2011 Yuan.

Table B2: PCA Weights (Component Loadings).

Cognitive Index	
Variable	Loading
Immediate Word Recall	0.595
Delayed Word Recall	0.588
Serial 7	0.414
Self-Reported Memory	0.137
Knows Current Month (Yes=1)	0.331

Table B3: ITT and LATE Estimates on Cognition using Propensity Score for NRPS Participation.

	Immediate Word Recall[a]	Delay Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
	NRPS (=1 if Propensity >= Mean + .5 SD			
Offered NRPS × Above60[c]	-0.052	-0.182***	-0.234**	-0.054
	(0.053)	(0.060)	(0.103)	(0.042)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.238	0.220	0.254	0.317
Observations	18,487	18,487	18,487	18,487
	NRPS (=1 if Propensity >= Mean + 1 SD			
Offered NRPS × Above60[c]	-0.075	-0.198***	-0.273***	
	(0.052)	(0.059)	(0.101)	-0.079* (0.041)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.230	0.216	0.247	0.310
Observations	20,309	20,309	20,309	20,309
Panel B (TOT):				
	NRPS (=1 if Propensity >= Mean + .5 SD			
PrNRPS × Above60[d]	-0.112	-0.396***	-0.508**	-0.118
	(0.116)	(0.133)	(0.226)	(0.091)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	458.733	458.733	458.733	458.733
Observations	18,487	18,487	18,487	18,487
	NRPS (=1 if Propensity >= Mean + 1 SD			
PrNRPS × Above60[d]	-0.161	-0.427***	-0.588***	-0.170*
	(0.112)	(0.129)	(0.220)	(0.088)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	348.111	348.111	348.111	348.111
Observations	20,309	20,309	20,309	20,309

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory and orientation. (c) Our DDD coefficient: Policy instrument interacted with an indicator for being over 60 years old. (d) Individual-level participation variable constructed from propensity score is instrumented with the policy instrument. Individual level controls: Above60 (1= Yes), Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community*Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community*Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table B4: ITT and LATE Estimates on Cognition Omitting Particular Communities.

	Immediate Word Recall[a]	Delay Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
	Sample excluding communities with less than 4 participants			
Offered NRPS × Above60[c]	-0.093*	-0.194***	-0.287***	-0.091**
	(0.054)	(0.058)	(0.101)	(0.040)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.226	0.213	0.245	0.310
Observations	19,566	19,566	19,566	19,566
	Sample excluding communities with less than 7 participants			
Offered NRPS × Above60[c]	-0.100*	-0.198***	-0.297***	-0.089**
	(0.054)	(0.058)	(0.102)	(0.041)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.225	0.211	0.243	0.307
Observations	19,057	19,057	19,057	19,057
Panel B (TOT):				
	Sample excluding communities with less than 4 participants			
NRPS Participation × Above60[d]	-0.228*	-0.472***	-0.700***	-0.221**
	(0.130)	(0.141)	(0.246)	(0.098)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	243.0807	243.0807	243.0807	243.0807
Observations	19,566	19,566	19,566	19,566
	Sample excluding communities with less than 7 participants			
NRPS Participation × Above60[d]	-0.238*	-0.472***	-0.711***	-0.213**
	(0.130)	(0.139)	(0.243)	(0.098)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	260.183	260.183	260.183	260.183
Observations	19,057	19,057	19,057	19,057

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. Individual level controls: Above60 (1= Yes), Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community×Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level

Table B5: ITT and LATE Estimates on Direct Measures of Health Varying the Definition of Instrument.

	Immediate Word Recall[a]	Delay Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
	Offer NRPS (=1 if at least 4 in community participate)			
Offered NRPS × Above60[c]	-0.070	-0.198***	-0.268***	-0.088**
	(0.049)	(0.057)	(0.096)	(0.039)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.229	0.216	0.247	0.313
Observations	21,202	21,202	21,202	21,202
	Offer NRPS (=1 if at least 7 in community participate)			
Offered NRPS × Above60[c]	-0.061	-0.178***	-0.239**	-0.078**
	(0.049)	(0.057)	(0.096)	(0.039)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.229	0.215	0.247	0.313
Observations	21,202	21,202	21,202	21,202
Panel B (TOT):				
	Offer NRPS (=1 if at least 4 in community participate)			
NRPS Participation ×	-0.162	-0.458***	-0.620***	-0.204**
Above60[d]	(0.115)	(0.135)	(0.226)	(0.091)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	240.613	240.613	240.613	240.613
Observations	21,202	21,202	21,202	21,202
	Offer NRPS (=1 if at least 7 in community participate)			
NRPS Participation ×	-0.140	-0.405***	-0.544**	-0.176**
Above60[d]	(0.113)	(0.132)	(0.222)	(0.089)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	258.336	258.336	258.336	258.336
Observations	21,202	21,202	21,202	21,202

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. Individual level controls: Above60 (1= Yes), Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community×Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

Table B6: ITT and LATE Estimates on Cognition. City-Level Analysis (Heilongjiang).

	Immediate Word Recall[a]	Delay Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
		CHARLS Data		
Offered NRPS × Above60[c]	-0.317	-0.151	-0.468	-0.213
	(0.267)	(0.263)	(0.476)	(0.189)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.113	0.080	0.106	0.149
Observations	178	178	178	178
		Admin Data		
Offered NRPS × Above60[c]	0.272	-0.288	-0.016	-0.057
	(0.693)	(0.762)	(1.308)	(0.507)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	.107	.079	.099	.145
Observations	178	178	178	178
Panel B (TOT):				
		CHARLS Data		
NRPS Participation ×	-1.558	-0.741	-2.299	-1.047
Above60[d]	(1.374)	(1.278)	(2.369)	(0.966)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	10.075	10.075	10.075	10.075
Observations	178	178	178	178
		Admin Data		
NRPS Participation ×	-0.557	0.590	0.033	0.117
Above60[d]	(1.407)	(1.595)	(2.682)	(1.044)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	8.526	8.526	8.526	8.526
Observations	178	178	178	178

Notes: The table reports estimates of the DDD estimator for the NRPS treatment effect. (a) Word recall tests: Immediate Recall = [0,10], Delayed Recall = [0,10] and Total Recall = [0,20]. (b) We created the Cognitive Memory Index using principal component analysis, combing measures of short/long term memory, working memory (Serial -7 Test) and orientation (Knowing the Current Month), and self-rated memory. (c) DDD coefficient: NRPS availability interacted with an indicator for being over 60 years old. Individual level controls: Above60 (1= Yes), Marital Status (=1 if Married), Gender (=1 if Female), Education Levels (Base Group is illiterate with no formal education), # of Household Residents. Columns 1-4 are estimated using Ordinary Least Squares (OLS) with Community and Year FE. Panel A is estimated using Ordinary Least Squares (OLS) with Community, Year and Community×Year FE. Panel B is estimated using Two-Stage Least Squares (2SLS) with Community, Year and Community×Year FE. Clustered standard errors at the community level reported in parenthesis.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.