Pension Benefits, Early Retirement and Human Capital Depreciation in Late Adulthood*

Plamen Nikolov[abcd☆] and Alan Adelman[a]

Abstract. Historically, economists have mainly focused on human capital accumulation and considerably less so 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 the introduction of a retirement program in China, we examine how the introduction of a retirement program influences individual cognition. We find large negative effects of pension benefits on cognitive functioning among the elderly. We detect the most substantial impact of the program to be on delayed recall, which is a significant predictor of the onset of dementia. We show suggestive evidence that the program leads to larger negative impacts among women. We show that retirement plays a significant role in explaining cognitive decline at older ages.

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

---

*We thank Charlotte Williams, Matthew Bonci, Sharon Itin, Hongyi Duan, Xu Wang and Jian Deng for outstanding research support with this project. We thank David Cutler, Jonathan Kolstad, Subal Kumbhakar, Eric Edmonds, Daniel Millimet, Carly Urban, Zoe McLaren, Carly Trachtman, 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.

☆Corresponding Author: Plamen Nikolov, Department of Economics, State University of New York (Binghamton), Department of Economics, 4400 Vestal Parkway East, Binghamton, NY 13902, USA. Email: pnikolov@post.harvard.edu

[a] State University of New York (Binghamton)
[b] IZA Institute of Labor Economics
[c] Harvard Institute for Quantitative Social Science
[d] Global Labor Organization

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 so on the causes and consequences of human capital depreciation. Recent evidence from neuropsychology, however, suggests that the human brain is malleable and open to enhancement even in late adulthood. In this paper, we examine the effect of the introduction 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 (Holzman, Robalino, and Takayama 2009).[1,2] We us the rollout of the pension program to examine how program participation and early retirement affect human capital depreciation and cognitive performance. Cognitive aging is an important and complex phenomenon, and its economic or policy causes are not well understood.

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 who have contributed towards the NRPS can start receiving a basic pension from the government and a portion from the individual's contributions to their account balance. We examine the effect of the pension program on cognitive aging and specifically on the link between early retirement, pension benefits, and cognitive decline in old age.

[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 (Musalem and Ortiz 2011). The program was financed from comingled funds: the first source of the program cost was financed by the local and federal funds; the second source came from individual contributions. The central government subsidizes 100 percent of the program cost in provinces with low fiscal capacity, whereas the federal subsidies constitute only 50 percent of the total funding in wealthier provinces with high fiscal capacity.

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 an individual's ability to process information. Elderly individuals are increasingly required to make complex financial, health, and long-term-care decisions, with consequences for their health and welfare.[3] Understanding the causes of cognitive decline is also crucial for policy—the relationship between cognitive aging and productivity matters for 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 any 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 triple-difference (DDD) estimator[4], or a difference-in-difference-in-differences methodology, to identify the causal effect of NRPS participation 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 beneficiaries 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)—that is nationally representative of individuals ages 45 and above. The CHARLS, a sister survey of the U.S.-based Health and Retirement Survey, directly tests cognition with a focus on two critical cognitive domains: *episodic memory*, and components of *intact mental status*. Episodic memory captures aspects of

[3] Cognitive ability remains an important factor associated with healthy aging (Rowe and Kahn 1997). 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 Ravallion et al. (2005) and Gruber (1994).
[5] Because our primary data source refers to these administrative units as communities, we refer to these municipalities as communities from now on.

fluid intelligence, whereas the mental intactness measure captures both fluid and crystallized intelligence.

Our analysis yields several interesting results. First, the NRPS program has a significantly negative effect on cognition among individuals aged 60 or above. Retirement programs are introduced and geared towards ensuring the welfare of aging adults (Cutler and Johnson 2004). Nevertheless, we provide strong evidence for a clear case of how the introduction of a pension program leads to 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. For the total word recall outcome, relative to the cognitive performance on tests before the provision of program benefits, the estimated effect size for cognitive decline, associated with approximately four years of program exposure, is 12 percent of a standard deviation (or approximately five percent of the average baseline score on the cognition measure). We can benchmark the effect size of our findings to general ability measures: a 5-percent drop in the average total recall score, due to NRPS participation, is approximately equivalent to a decline in general intelligence by 1.7 percent (relative to the general population).[6] Our main results are robust to several different specifications that especially test the validity of the proxy measure of NRPS participation. Furthermore, we examine if longer retirement duration 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 as well as

[6] Ackerman et al. (2005) and Healey et al. (2014) investigate the relationship between recall memory measures and general intelligence (fluid intelligence). Based on a meta-analysis, 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.

generous government subsidies for particular contribution levels could end up boosting permanent income for some people. Consequently, this income boost could reduce incentives for full labor market participation on 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, the reduction of 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 on 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 program's 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 the large magnitude of the estimated effect sizes based on our analysis 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*.[7,8,9]. 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 the results found in these two studies, our estimates are also negative; however, they are considerably lower than the ones estimated by Rohwedder and Willis (2010) and Mazzonna and Peracchi (2012).[10]

Neurological research demonstrates that proxy measures of delayed recall memory are highly accurate detectors of dementia (Welsh et al. 1991).[11] 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[12], 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 in at least four significant ways. First, this study is the first, to the best of our knowledge, to examine the effects of pension participation on

[7] 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.[7] Conversely, Coe et al. (2012) find no conclusive evidence with data from the US Health and Retirement Survey (HRS).

[8] Other recent empirical studies also examine the effect of the NRPS on other individual or household-related outcomes. For example, Nikolov and Adelman (2009ab) examine the effects of the NRPS program 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.

[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 at al. 2012, Mazzonna and Peracchi 2012, Bingley and Martinello 2013, de Grip et al. 2015). Additionally, we find that the provision of pension benefits 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 the symptoms of 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] 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 greater incidence of dementia due to pension program participation.

[12] If gender differences in cognitive decline exists, this fact could have alarming implications for pension policy. The average performance on cognition tests for Chinese females is much lower than the performance of Chinese males; the gender difference is particularly pronounced among older Chinese cohorts (Lei et al. 2012). Coupled with the fact that females have a longer life expectancy (Liu et al. 2009), 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 (Wang 1993).

individual cognitive performance in the context of a developing country.[13,14] China is a particularly suitable context to examine this issue, given the country's population size, and the growing share of its elderly population.[15] Second, we illuminate how program participation affects a 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 one's life—for example, financial planning for retirement (Banks and Oldfield 2007) and medical treatment adherence to planning for sequential activities (Fillenbaum et al. 1988). In this paper, we focus on proxies of cognition, specifically episodic memory, which neurobiology research documents to be particularly sensitive to the aging process. Several studies highlight that this domain is the first to exhibit setbacks as aging sets in (Souchay et al. 2000; Tulving and Craik 2000; Prull et al. 2000). The second reason relates to its provision of high individual variation across individuals as opposed to other cognitive measures. [16] 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

[13] 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 measures of health, including mobility, self-care, usual activities, and vision. In this study, we show 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 that determine cognitive depreciation likely differ 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.

[14] 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 the after the introduction of 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.

[15] China's population is aging rapidly. In 2007, approximately 11 percent of China's population was aged 60 or over, making up 21 percent of the world's elderly population (UN 2007). Our analysis focuses on China, the country with the largest population in the world, home to 1.4 billion people. 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 dynamics resemble the economies of low-income countries. Therefore, our findings have important implications for other low-income countries.

[16] For example, the word learning and recall tasks do not exhibit floor or ceiling effects (excess of maximum or minimum values), and 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, Ludwig, and Liebman (2004) and 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.

high- and middle-income countries.[17] The survey harmonization of cognition measures across surveys can enable additional analyses and international comparisons across different country settings based on data from these retirement surveys.[18] Finally, we provide suggestive evidence on the underlying mechanisms leading to cognitive decline among the elderly.

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.[19] 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.

[17] 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 core component of the survey 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. Because of its successful implementation, the survey has become a model around the world for similar surveys that specifically examine issues of health and retirement. Currently, harmonized constructs on health and demographic information exists across 18longitudinal 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, etc.) 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.

[18] 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).

[19] Vilela (2013) reviews the history and the evolution of China's pension policy since the establishment of 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 a stronger emphasis on universal coverage of formal and informal workers alike. Vilela (2013) highlights three distinct historical phases of the country with regard to 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).

In addition to the crumbling fiscal sustainability of the decentralized pension programs, the demographics of the country 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. Due to the combination of considerably lower fertility as a result of the one-child policy and reduced mortality, the population structure in the country shifted towards older age groups, resulting in a rapidly aging population.[20] Chinese government policy was necessary to tackle the growing demographic challenges and old-age poverty in rural areas. The government introduced a rural pension program in 1986 by piloting the rollout among rural residents. However, a combination of poor governance and additional financial challenges halted the expansion of this program throughout the 1990s.[21]

In the early 2000s, the rural pension system faced continuing challenges related to financial sustainability. The government, under the new Hu-Wen administration that assumed leadership in 2003, ostensibly adopted a reform-oriented approach for the country's social protection system. 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.

[20] The growth rate of the population 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 age-sex distribution of the population. However, from 1964 to 1982, the oldest age groups did experience a considerable proportional increase while the percentage distribution of the two youngest age groups declined substantially.
[21] 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 social basic 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 expansion of the rural pension program and it was substantially scaled back by 1999 and the early years of the 2000s. Pension coverage under the system declined from 80.25 million participants in 1998 (approximately 11 percent of the total rural population) to mid-50 million in 2007.

In 2009, China launched the NRPS. Participation in the NRPS was available to all rural residents over the age of 16 years, 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 used in mainland China. 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 the age of 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.[22] Participants between the ages of 45 and 60, with fewer than 15 years of contributions, were encouraged to increase their monthly payments so they could cover the absence of prior contributions before age 45.

The NRPS aimed to achieve full geographic coverage in rural parts of China 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). By the end of 2011, over 50 percent of rural residents contributed to the NRPS by , and total participation in the program grew from 87 million to 326 million people from 2009 to the end of 2011 (Quan 2012).[23]

[Figure 1 about here]

Three major factors account for the NRPS expansion between 2009 to 2011. First, the country's high economic growth rate played a considerable role. Between 2009 and 2011,

[22] 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).

[23] We examine whether the age characteristics of communities that implemented the NRPS earlier differed from the age characteristics of communities that adopted the program later and we find no evidence of differences in the average age. It is possible that the participating and non-participating areas differed on other socio-economic characteristics; we address this issue in Section III.

China's economy grew at an average annual rate of 9.3 percent, which provided robust fiscal capacity for the rollout of the massive social protection program. Second, because of increasing income inequality and demographic pressures, demand for the program was substantial. Third, pension reform and the expansion of the pension program into the country's rural areas were fundamental political priorities for the Hu-Wen administration.

Program Eligibility and Benefits. The total amount available for program payouts is determined by individual contributions and matching funds by local governments. Local governments must match at least 30 RMB annually per individual contribution.[24] 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. [25,26]The pension 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.

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

[24] 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.

[25] A participant may stop contributing for a few years and make up for the missed contributions later; they would only lose the subsidies for the years that 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.

[26] 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. If individuals are to participate, the best investment strategy (in terms of maximizing the net present value of the NRPS contributions) is choosing the lowest premium level and start contributing as late as possible. The study shows that given a local subsidy of 30 RMB per person per year, only the option of contributing for less than 21 years at the lowest premium level (100 RMB) obtains positive net benefit (under additional assumptions about the annual interest rate, the timing regarding when the pension benefits are claimed, and the annuity factor). The rates of return on contributions above the lowest contribution level are limited to the rate of return on pension assets, which is specified as the one-year bank deposit rate. Therefore, there is a strong incentive for individuals to choose 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.

the individual account funds, an individual can typically 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 the country's 2009 rural annual per capita net income. Based on the chosen contribution level, a government subsidy is then 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. The sum of the individual contributions and government subsidies are deposited in the individual's account. The interest rate is the one-year base rate, according to the People's Bank of China (the Central Bank of China), 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 is compounded yearly.

Individuals who are aged 60 or older can start receiving the basic pension every month without making any contributions if all of their eligible children living in the same village participate in the program. 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.[27] 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.

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

The monetary benefits paid out to participants follow the "*139 Rule*", which is a number based on the average life expectancy (in months) at age 60.[28] The rule follows a basic formula for the calculation of 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 empirical analysis draws on data obtained from the CHARLS. Due to the survey's rich data on various proxies of cognition, we draw on this source for our primary analysis in which we compare the cognitive outcomes of individuals with access to NRPS program benefits. The survey also directly captures a person's participation in the NRPS program. The second part of our analysis draws on data from the China Health and Nutrition Survey. We use this second survey for our analysis because it collects data for years before the implementation of the NRPS. We also use the survey to conduct several additional robustness checks, which we describe in detail 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. The CHARLS provides data on demographic characteristics, family structure, cognition, health, pension and retirement, work, household wealth, income, and consumption. The timing of the CHARLS is ideal for our analysis since it was conducted approximately a year after the NRPS implementation.

[28] The individual account has a rate of return that is set to the People's Bank of China's one-year deposit rate. The "139" Rule was adopted from the already established Urban Pension Scheme.

Our analysis sample, based on an individual-level panel, consists of 15,990 individuals across two waves from 429 communities in 121 cities across 28 provinces. 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] CHARLS directly collects information on individual participation in various government programs, including the NRPS. For our primary analysis sample, we drop observations with an urban *Hukou* status because individuals who are attached to an urban *Hukou* are ineligible to participate in the NRPS but instead participate in urban pension schemes.[31] 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 extensively 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

[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.

[31] The Urban Social Pension Scheme was established in 2011 and rapidly expanded tot 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 the several important 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 choice for 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 continues to answer unrelated questions for several minutes until prompted to recall the original word list. This provides the delayed recall score, ranging from zero to ten.

second cognition measure tests one's 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 exhibit setbacks as aging sets in (Souchay et al. 2000; Tulving and Craik 2000; Prull et al. 2000).[35,36] In addition to several cognitive tests, respondents are asked to rate their memory based on a 5-point scale. Based on this scale, we create a binary indicator to denote if an individual is in good health.

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 over time into a stock of human capital. Our first measure, episodic memory, captures aspects of fluid intelligence as it encompasses people's 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 people's ability to infer and to 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

[34] This includes recognition of date: (month, day, year, season (CHARLS allows use of the lunar calendar in addition to the Gregorian calendar), day of the week), how the respondent rates their own 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. Based on these measures, we compute the sum of two scores—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) and 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 is as follows: (1) Excellent (2) Very Good (3) Good (4) Fair (5) Poor.

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. About three-quarters of the sample work in agriculture: 72 percent among participants and non-participants alike. The rural sample reported low levels of educational attainment—approximately 46 percent to 48 percent report having completed at least a secondary school degree. In terms of health, approximately one-quarter, 27 percent of participants and 26 percent of non-participants, report being in "poor/fair" health status.

[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 prior to 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

[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. By design, the first component, based on the largest eigenvalue, contains the most information, whereas the last component contains the least. We reduce the set of original cognition proxies into the 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 B1 reports the index component loadings based on the survey's cognitive measures.

[40] Participants scored slightly higher on the word recall tests. The average score on 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 it was for non-participants, 2.91 and 2.89 respectively. Approximately 84 percent of participants and non-participants correctly named the current month. The cognitive memory index, based on the PCA, exhibits a higher average for participants than it does 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.

collect data on various economic and social determinants of individual health and nutritional status (UNC-Chapel Hill 2010). Adults aged 55 and older were asked to provide a record of their daily living activities and were given various cognition tests. The CHNS also collected information on proxy measures of memory and cognition, very similar to those collected by the CHARLS. The CHNS adopted similar cognitive screening items because its cognition proxies were also based on 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]

III. Empirical Strategy

Our primary identification strategy relies on across-municipality variation in the NRPS implementation. In particular, we exploit a source of identifying variation due to the program's staggered rollout, between 2011 and 2013, across the rural parts of the country. The DDD analysis exploits the rollout of the program 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 who live in communities that adopted the NRPS.

[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 better 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.

A. Estimating Equations and Triple Difference Estimation

Given the staggered rollout of the program at the community level, our identification strategy relies on the timing differences across communities regarding the adoption of the pension policy.[45] 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) that are 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*). The linkage between the various administrative layers within the CHARLS allows us to link each community identifier within the survey with a person's place of living and his/her response to whether they participate in various government programs, including the NRPS. This process allows us to define the variable $OfferNRPS_{ct}$ based on responses from the individual-level data.[46] We examine the impact of the NRPS provision on cognition using the following specification[47]:

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

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

[45] Demographic information in the CHARLS is only available at the community (*shequ*) level.

[46] 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.

[47] 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., Abowd et. al. 1999, Andrews et. al. 2006, Duflo 2004).

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, education, gender, age, age squared, household size, and marital status. ϕ_c and μ_t are community-level and time fixed effects, respectively. Community-level fixed effects allow us to control for time-invariant characteristics that affect the likelihood that a community implements the NRPS. The time fixed effects account for communities-specific features that do not change over time.[48,49] We use community-time fixed effects, $\phi_c \times \mu_t$, to control for community differences during the implementation of the NRPS. The DDD design is the most appropriate choice, as it controls for potential region-specific effects and is based on a similar policy rollout in other empirical studies, such as Katz (1996), Gruber (1994), and Rossin (2011).

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, our coefficient of interest. The *common trends* identification assumption, which underlies the DDD design in our estimation, posits that important factors that influence the study outcomes are either time-invariant group attributes or time-varying factors that are group invariant. In summary, the identification assumption implies that communities that adopted the NRPS program would otherwise have changed in a manner similar, on average, to the communities that did not adopt the NRPS. To check whether the triple difference is an appropriate strategy to examine the effect of the NRPS program, we test the common trends assumption for the pre-

[48] We cluster the standard errors by community and age groups based on Bertrand, Duflo, and Mullainathan (2003).
[49] In Online Appendix B, we report additional robustness checks in which we cluster the standard errors by community and age. Our results are robust to community and age-specific clusters.

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 launch of the NRPS program 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 this particular 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 proceed with testing the common trends assumption at the province level using the CHNS data 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 is based on two important features. First, it is based on the available 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, our treatment definition of a province is defined continuously as the "treatment intensity" of a given province. Based on the continuous variable, we then code the province's treatment status with a binary variable. The province's treatment status is set to one if more than a given threshold of communities reported participating in the NRPS, and zero otherwise. Based on this reconstructed definition of a "treatment status" for a province, and only for this formal test of the common trends' assumption, we then proceed to use data from

the CHNS. We analyze data from the CHNS before the NRPS's introduction to test for any potential common trends between treated and control provinces. 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.[50,51] 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 the 2004, 2006, and 2009 waves, we can estimate the following specification:

$$Y_{ict} = \beta_0 + \beta_{-3} D_{ict-3} + \beta_{-1} D_{ict-1} + \phi_c + \mu_t + \phi_c \times \mu_t + \varepsilon_{ict}\,, \tag{2}$$

where Y_{ict} is the cognition proxy, and ϕ_c, μ_t, and $\phi_c \times \mu_t$ are community, time and community-time fixed effects, respectively. Because of the triple-difference estimation, we include the triple interactions $D_{ict} = (OfferNRPS_{ct} \times Above60_{ict})$ for the first and last pre-treatment periods a.[52] D_{ict} is defined in the same way as in our main triple-difference specification. 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 outcomes between the treatment and control areas are the same.[53] Therefore, this exercise provides no empirical evidence to undermine the validity of the common trends assumption.

[50] 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 information collected from the 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 the treatment status for each province and we re-estimate our specifications.

[51] 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.

[52] In this specification, the second pre-treatment period is omitted.

[53] 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 main 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 not only statistically insignificant but also unstable across different threshold specifications, a fact that further undermines any evidence of robust differences, at the province level, in pre-trends.

B. Instrumental Variable Estimation

To address the possibility of endogenous individual participation, we augment the DDD analysis by instrumenting individual program participation with the program availability at the community level. Specifically, we re-estimate specification (1). However, we also use $OfferNRPS_{ct}$ to instrument for a person's participation in the NRPS, an instrumented difference-in-differences design as in Abdulkadiroglu et al. (2011), Hudson, Hull, and Liebersohn (2017), and Nunn and Qian (2014). We code $OfferNRPS_{ct}$ as a binary variable, and 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:

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

$\widehat{NRPS}_{ict}$ represents individual enrollment in NRPS, and we instrument it with $OfferNRPS_{ct}$. $\boldsymbol{X}_{ict}$ is a vector of individual-level controls, and 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 program's impacts on cognitive outcomes based on specifications (1) and (3). Table 2 reports the results. Columns 1 through 4 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, individuals in program areas score worse by approximately 12 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 provision of the NRPS benefits leads to a 0.10-point reduction in the composite score (equivalent to about 7 percent of one standard deviation) for the intent-to-treat specifications. Based on the 2SLS estimation, the effect size associated with the effect of 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. The associated effect size for this particular outcome 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 program provision 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 size of the effect size estimates based on the ITT specification.[54] 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

[54] 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.

for all cognition measures: the delayed word recall task, total word recall score, and composite memory index.

The analysis so far focuses on data from the CHARLS, based either 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 created a reconstructed definition of retirement.[55] 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.

[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 influence the program exerts 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 particular dataset. Program participants who retire exhibit decline (i.e., negative coefficients reported in Columns 1 through 3) in their performance on the cognition tests. Among people who retire, we find that the effect size associated with the composite memory measure is sizable, negative, and statistically significant at the five percent level.

[55] 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 work days missed for health reasons has been more than 300 per year for three consecutive survey waves, reported year of retirement is prior to 2009, and the survey respondent indicated that the formal retirement is processed.

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, increased exposure to the program benefits should lead to a larger effect size of cognitive decline. 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 2.9 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 as compared to 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.[56]

[Table 4 about here]

B. Mechanisms

Obtaining access to the retirement program benefits likely exerts changes on a whole 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 that are likely also affected by NRPS.

[56] 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.

Specifically, we examine how NRPS participation affected four major groups of activities of program beneficiaries: (1) labor market activities, (2) mental stimulation activities, (3) social engagement, and (4) various health behaviors, time use, and health care utilization. If we observe no change in a measure that could play a mediating role in influencing cognition, then we would have a strong indication that the causal pathway does not operate via that mediating factor (or group of factors). Table 5 reports 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 14 in Table 5). 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 activities associated with mental stimulation and social engagement. Specifically, NRPS participants report a substantial decrease in their engagement in adult education and board games. They also report lower social community engagement or volunteering as compared to the comparable individuals who did not obtain NRPS benefits.

[Table 5 about here]

These results warrant an explanation of the net effects on cognition. On the one hand, NRPS participation leads to numerous benefits on an individual's 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 so 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 6 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 6 about here]

The results reported in Table 6 echo the negative effect of the NRPS program on cognitive measures. Table 6 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). Comparing the effect size estimates for the aggregated measure reported in Column 4, we show that the NRPS effect on cognitive decline among females is approximately double the effect as compared to the cognitive decline among males. However, the coefficient for the DDDD estimator is not significant at conventional levels for statistical significance.[57]

V. Robustness Checks

In this section, we perform a series of exercises aimed to re-examine the validity of our empirical approach. 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 are not eligible to 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.

The main results survive these extension exercises, a fact implying that our main results are unlikely to be driven by secular trends, alternative contemporaneous policies in the same areas where NRPS was implemented, or by unobserved shocks to our 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 individuals who are neither eligible for NRPS participation nor its program benefits. Therefore, in theory, when

[57] We examine for 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 labor market, particularly for females in self-employment.

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 elderly individuals without children who live in rural administrative districts (rural *Hukou*) because these individuals are ineligible for the NRPS. However, for this particular falsification exercise, we reconstruct the analysis sample and employ the opposite approach. Only for this particular exercise, we include a sample of respondents who are ineligible for NRPS participation. This sample comprises pensioners in an urban pension program or elderly individuals (aged 60 and above) who did not contribute to the NRPS before reaching age 60 and reported no current children who live 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) those who are elderly and without children who happen to live in rural administrative districts. The main objective of this falsification exercise is to examine whether the effect of the NRPS on the cognition outcomes for this "placebo" sample would differ between individuals who live in areas that offered NRPS coverage and individuals who live in areas that did not. If specifications (1) and (3) yield no spurious results, this specific falsification exercise should produce non-significant coefficient estimates for the coefficients associated with the NRPS effect on cognition outcomes.[58] We re-estimate specifications (1) and (2), as described in the main empirical approach, based on the

[58] 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 but they themselves are not eligible (nor do they receive program benefits) for the NRPS. As we show in Section III, NRPS program beneficiaries considerably lower their social interactions, so if spillovers are plausible via changes in social interaction (one viable mechanism for social spillovers) then our analysis will pick up some effect among non-beneficiaries who live in areas that offer NRPS and we do not detect such changes.

placebo sample. In Appendix A, Table A3, we report the results based on this falsification exercise.

Table A3 reports non-significant estimates for the effect of the NRPS program on the various 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.[59]

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 administrative communities' incorrect reports of NRPS program implementation within their boundaries. 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 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

[59] 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 female household members, the number of daughters in the household, and the mother's educational level. 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.

likelihood that an individual participates in the NRPS. We define the NRPS participation status based on a propensity score matching approach. Specifically, we predict the NRPS participation status (at baseline) based on a combination of the respondent's 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 $\mathrm{PrNRPS_{ic}}$. 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 $\widehat{\mathrm{NRPS}}_{\mathrm{ic}}$.

Next, using this redefined measure of NRPS participation based on the propensity score approach, we proceed with the DDD analysis as in the main portion of our results section based on specifications (1) and (3). The NRPS program participation in this analysis uses the reconstructed variable ($PrNRPS_{ic}$), which is 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).[60] We report the results based on this alternative definition of program participation in Online Appendix Table B2. The reported results demonstrate that the pattern of our findings is indeed robust to this alternative definition of NRPS program participation.

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

[60] 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 approach relies on an 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).[61]

Therefore, we verify the robustness of our approach with an alternative definition of the variable $OfferNRPS_{ct}$. Specifically, the objective of this additional exercise is to correct for possible contamination of what communities are coded as treated (or indicating implementation of the NRPS program). The source of this measurement could be due to false reporting at the individual level. To tackle this potential issue, 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 for when the community indicator switches from zero (non-participating) to 1 (participating) based on an even higher threshold, of least seven individuals within the community, for each community in the CHARLS, reporting participating in the NRPS.

We report the results from these additional analyses in Online Appendix Table B3. The results demonstrate that our original estimates are robust to the alternative and much more conservative definitions of the threshold, which determines when the variable $OfferNRPS_{ct}$ switches its binary values.

[61] 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 main analysis, we define the implementation of the NRPS program 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, then we define the community as one having implemented the NRPS. However, it is possible that communities with a very small number of NRPS participants are systematically related to a set of other factors that we cannot observe in our survey data and account in our analysis. 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 B4 reports the results for this extension exercise. The pattern of the results remains consistent with the main results.

Using Online Administrative Data. Using data from online sources, we perform a final robustness check to bolster the direction of the estimated program effects based on the main analysis. For this empirical exercise, we comb data available from Chinese newspapers (online or in paper format) based on public announcements regarding geographic participation in the 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.[62] 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, which is the analysis period in our main estimation approach. 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 the definition of the *community* unit between the CHARLS and the actual administrative units from the public announcements regarding NRPS implementation. We are unable to map the actual communities (within cities) to the communities (or the variable *community I.D.*) in the CHARLS survey, which is the analysis unit in our main analyses; this is because the definition of a community in the CHARLS differs from the definition of an administrative *community unit* (available in online records). Although we are unable to re-estimate both the main analysis (which we perform at the community level) and replace the survey data from the CHARLS with administrative data for community-level participation, we can rerun our previous specifications at the city level (i.e., at a higher

[62] In this additional estimation exercise, we can estimate the original specifications at the city level. Specifically, we can compute the treatment intensity (for city participation) as follows: $city\_participation_t$=(# communities in a city that offered the$NRPS_{ct}$) (# total communities in a city)$_{ct}$. The main advantage of this additional robustness check is that we are able to 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 constant. A major disadvantage of this approach is that we are able to re-estimate the specifications from Section III only at the city level (and only for the Heilongjiang Province, for which we are able to obtain data on city or community announcements regarding NRPS implementation). This implies that in this additional analysis the number of number of observations is low, which limits the statistical power for statistical inference.

geographic level than the community level at which we perform our primary analysis). We can conduct this auxiliary analysis at the city level for two reasons: (1) we can observe the number of communities within each city that definitively implemented the NRPS program based on the online public announcements, and (2) we can observe the total number of all communities within each city; the total number of these communities is a fixed number.

Therefore, we redefine our main treatment. Specifically, and only for this additional empirical exercise, we change the definition of the treatment variable 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).[63] Based on this reconstructed definition of the treatment variable, we re-estimate this additional robustness check at the city level as opposed to the community level, an estimation approach performed in the main analysis.

Using data for Heilongjiang Province, we re-estimate the empirical specifications outlined in Section III. However, we rely on a continuous definition of the main treatment variable. We report the results for this final robustness check at the city level in Online Appendix Table B5. This additional analysis (at the city level) relies on a minimal sample. However, despite this statistical power limitation, the pattern of the results echoes the pattern reported in the main analysis. Both the effect size and the direction of the program effects are consistent with the direction of the main estimates based on survey data from the CHARLS. The results based on these additional analyses provide additional evidence regarding the direction of the program impacts reported in our main analyses.

[63] The Heilongjiang province is ideal for this empirical exercise for a number of 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 main empirical approach relies on identifying variation based on both time and space. Third, most of the city-level implementation of the NRPS in this province occurred around 2013. Other cities or areas within provinces had either already adopted the NRPS program prior to 2013, or information on city-level NRPS implementation by year was not readily available online or via posted public records.

VI. Concluding Remarks

In this paper, we investigate the effect of China's NRPS program 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. Albeit surprising, 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; 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 the normal aging among the elderly and individuals more likely to witness an earlier onset of dementia in adulthood.

Furthermore, our findings support the *mental retirement* hypothesis that decreased mental activity results in atrophy of cognitive skills and suggests that retirement plays a significant role in explaining cognitive decline at an older age. However, further studies would be necessary to specify the effect of professional activities on cognition and, in particular, on other cognitive domains. Specifically, two additional areas will be of particular 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 the speed of individual cognitive decline? Second, it is essential to uncover and examine the underlying

mechanisms between a person's retirement and cognitive decline. A particularly 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 on 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

Abdulkadiroğlu, A., Angrist, J.D., Dynarski, S.M., Kane, T.J., Pathak, P.A., 2011. Accountability and Flexibility in Public Schools: Evidence from Boston's Charters And Pilots. Q J Econ 126, 699–748. https://doi.org/10.1093/qje/qjr017

Abowd, J.M., Kramarz, F., Margolis, D.N., 1999. High Wage Workers and High Wage Firms. Econometrica 67, 251–333. https://doi.org/10.1111/1468-0262.00020

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.

Adam, S., Bonsang, E., Germain, S., Perelman, S., 2007. Retirement and cognitive reserve: A stochastic frontier approach applied to survey data, University of Liège, CREPP Working Paper 2007/04. University of Liège, Liege, Belgium.

Andrews, M., Schank, T., Upward, R., 2006. Practical Fixed-Effects Estimation Methods for the Three-Way Error-Components Model. The Stata Journal 6, 461–481. https://doi.org/10.1177/1536867X0600600402

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

Banks, J., Oldfield, Z., 2007. Understanding pensions: Cognitive function, numerical ability and retirement saving. Fiscal Studies 28, 143–170. https://doi.org/10.1111/j.1475-5890.2007.00052.x

Bertrand, M., Duflo, E., Mullainathan, S., 2004. How much should we trust differences-in-differences estimates? The Quarterly Journal of Economics 119, 249–275. https://www.jstor.org/stable/25098683

Bingley, P., Martinello, A., 2013. Mental retirement and schooling. European Economic Review 63, 292–298. https://doi.org/10.1016/j.euroecorev.2013.01.004

Bloom, D.E., McKinnon, R., 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. Institute for the Study of Labor, Bonn, Germany.

Bonsang, E., Adam, S., Perelman, S., 2012. Does Retirement Affect Cognitive Functioning? Journal of Health Economics 31, 490–501. https://doi.org/10.1016/j.jhealeco.2012.03.005

Brandt, J., Spencer, M., Folstein, M., 1988. The telephone interview for cognitive status. Neuropsychiatry, Neuropsychology and Behavioral Neurology 1, 111–117.

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, F., Giles, J., O'Keefe, P., Wang, D., 2012. The elderly and old age support in rural China, 1st ed. World Bank Publications, Washington, DC.

Cheng, L., Liu, H., Zhang, Y., Zhao, Z., 2018. The health implications of social pensions: Evidence from China's new rural pension scheme. Journal of Comparative Economics 46, 53–77. https://doi.org/10.1016/j.jce.2016.12.002

China Center for Economic Research, 2013. Challenges of population aging in China. Evidence

from the national baseline survey of the China Health and Retirement Longitudinal Study (CHARLS) (No. Report 0916), Beijing: School of National Development, Peking University. Peking University, Beijing, China.

Coe, N.B., Gaudecker, H. von, Lindeboom, M., Maurer, J., 2012. The effect of retirement on cognitive functioning. Health Economics 21, 913–927. https://doi.org/10.1002/hec.1771

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

de Grip, A., Dupuy, A., Jolles, J., Boxtel, M. van, 2015. Retirement and cognitive development in the Netherlands: Are the retired really inactive? Economics & Human Biology 19, 157–169. https://doi.org/10.1016/j.ehb.2015.08.004

Dorfman, M.C., Holzmann, R., O'keefe, P., Wang, D., Sin, Y., Hinz, R., 2013. China's pension system: A vision, 1st ed. World Bank Publications, Washington, DC.

Duflo, E., 2004. The medium run effects of educational expansion: Evidence from a large school construction program in Indonesia. Journal of Development Economics 74, 163–197. https://doi.org/10.1016/j.jdeveco.2003.12.008

Eibich, P., 2015. Understanding the effect of retirement on health: mechanisms and heterogeneity. Journal of Health Economics 43, 1–12. https://doi.org/10.1016/j.jhealeco.2015.05.001

Feldstein, M., Liebman, J.B., 2002. Chapter 32: Social security, in: Auerbach, A.J.F., and Martin (Ed.), Handbook of Public Economics. Elsevier, Amsterdam, The Netherlands, pp. 2245–2324.

Fillenbaum, G.G., Hughes, D.C., Heyman, A., George, L.K., Blazer, D.G., 1988. Relationship of health and demographic characteristics to Mini-Mental State Examination score among community residents. Psychological Medicine 18, 719–726. https://doi.org/10.1017/s0033291700008412

Gruber, J., 1994. The Incidence of Mandated Maternity Benefits. The American Economic Review 84, 622–641. https://www.jstor.org/stable/2118071

Healey, M. Karl, Patrick Crutchley, and Michael J. Kahana. 2014. "Individual Differences in Memory Search and Their Relation to Intelligence." *Journal of Experimental Psychology. General* 143 (4): 1553–69. https://doi.org/10.1037/a0036306.

Heckman, J.J., 2000. Policies to foster human capital. Research in Economics 54, 3–56. https://doi.org/10.1006/reec.1999.0225

Holzman, R., Robalino, D., Takayama, N., 2009. Closing the Coverage Gap. The Role of Social Pensions and Other Retirement Income Transfers. World Bank, 1st ed. The World Bank, Washington, DC.

Hudson, S., Hull, P., Liebersohn, Jack, 2017. Interpreting Instrumented Difference-in-Difference (No. Massachusetts Institute of Technology Mimeo). Massachusetts Institute of Technology, Cambridge, MA.

Jones, L., Yilmazer, T., 2018. The Long-Run Effect of the Earned Income Tax Credit on Cognitive Ability, in: Summer Research Workshop. Presented at the Summer Research Workshop, University of Wisconsin, The Institute for Research on Poverty, University of Wisconsin, The Institute for Research on Poverty, Madison, Wisconsin.

Katz, L.F., 1996. Wage Subsidies for the Disadvantaged (No. NBER Working Paper No. 5679),

National Bureau of Economic Research, Mimeo. National Bureau of Economic Research, Mimeo, Cambridge, MA.

Kling, J.R., Liebman, J.B., Katz, L.F., 2007. Experimental Analysis of Neighborhood Effects. Econometrica 75, 83–119. https://doi.org/10.1111/j.1468-0262.2007.00733.x

Kling, J.R., Ludwig, J., Katz, L.F., 2005. Neighborhood Effects on Crime for Female and Male Youth: Evidence from a Randomized Housing Voucher Experiment. The Quarterly Journal of Economics 120, 87–130. https://doi.org/10.1162/0033553053327470

Korniotis, G.M., Kumar, A., 2011. Do Older Investors Make Better Investment Decisions? The review of economics and statistics 93, 244–265. https://doi.org/10.1162/REST_a_00053

Laakso, M.P., Hallikainen, M., Hänninen, T., Partanen, K., Soininen, H., 2000. Diagnosis of Alzheimer's Disease: MRI of the Hippocampus vs Delayed Recall. Neuropsychologia 38, 579–584. https://doi.org/10.1016/s0028-3932(99)00111-6

Langa, K.M., Ryan, L.H., McCammon, R.J., Jones, R.N., Manly, J.J., Levine, D.A., Sonnega, A., Farron, M., Weir, D.R., 2020. The Health and Retirement Study Harmonized Cognitive Assessment Protocol Project: Study Design and Methods. Neuroepidemiology 54, 64–74. https://doi.org/10.1159/000503004

Lei, X., Hu, Y., McArdle, J.J., Smith, J.P., Zhao, Y., 2012. Gender Differences in Cognition among Older Adults in China. The Journal of Human Resources 47, 951–971. https://www.jstor.org/stable/pdf/23798522.pdf

Lei, X., Zhang, C., Zhao, Y., 2013. Incentive Problems in China's New Rural Pension Program. Research in Labor Economics 37, 181–201. https://doi.org/10.1108/S0147-9121(2013)0000037010

Liu, J., Chen, G., Song, X., Chi, I., Zheng, X., 2009. Trends in Disability-Free Life Expectancy Among Chinese Older Adults. Journal of Aging and Health 21, 266–285. https://doi.org/10.1177/0898264308328978

Liu, T., Sun, L., 2016. Pension Reform in China. Journal of Aging & Social Policy 28, 15–28. https://doi.org/10.1080/08959420.2016.1111725.

Mazzonna, F., Peracchi, F., 2012. Ageing, cognitive abilities and retirement. European Economic Review 56, 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.

Meng, Q., Wang, H., Strauss, J., Langa, K.M., Chen, X., Wang, M., Qu, Q., Chen, W., Kuang, W., Zhang, N., 2019. Validation of neuropsychological tests for the China Health and Retirement Longitudinal study harmonized cognitive assessment protocol. International Psychogeriatrics 31, 1709–1719. https://doi.org/10.1017/S1041610219000693

Musalem, A.R., Ortiz, M.D., 2011. Governance and social security: Moving forward on the ISSA good governance guidelines. International Social Security Review 64, 9–37. https://doi.org/10.1111/j.1468-246X.2011.01409.x

Nikolov, P., Adelman, A., 2019a. Short-Run Health Consequences of Retirement and Pension Benefits: Evidence from China. Forum for Health Economics and Policy 21, 1–27. https://doi.org/10.1515/fhep-2017-0031

Nikolov, P., Adelman, A., 2019b. Do Private Household Transfers to the Elderly Respond to Public Pension Benefits? The Journal of the Economics of Ageing 14, Article 100204.

https://doi.org/10.1016/j.jeoa.2019.100204

Nunn, N., Qian, N., 2014. US Food Aid and Civil Conflict. Am.Econ.Rev. 104, 1630–1666. https://doi.org/10.1257/aer.104.6.1630

Ofstedal, M.B., Fisher, G.G., Herzog, A.R., 2005. Documentation of cognitive functioning measures in the Health and Retirement Study, University of Michigan, HRS Documentation Report DR-006. University of Michigan, Ann Arbor, MI.

Prull, M.W., Gabrieli, J.D., Bunge, S.A., 2000. Age-related changes in memory: A cognitive neuroscience perspective., in: Craik, Fergus.S., Timothy A. (Ed.), Handbook of Aging and Cognition. Lawrence Erlbaum Associates Publishers, Mahwah, NJ, pp. 91–153.

Qing, N.B., 2009. The useful trial to establish a new rural old-age insurance system: Investigation on the pilot project of Baoji, Shaanxi Province. Theoretical Horizon 6, 40–42.

Quan, L., 2012. Analysing the coverage gap in China, London, HelpAge International. HelpAge International, London, UK.

Ravallion, M., Galasso, E., Lazo, T., Philipp, E., 2005. What Can Ex-Participants Reveal about a Program's Impact? Journal of Human Resources 40, 208–230. https://doi.org/10.3368/jhr.XL.1.208

Rohwedder, S., Willis, R.J., 2010. Mental Retirement. The Journal of Economic Perspectives 24, 119–138. https://doi.org/10.1257/jep.24.1.119

Rossin, M., 2011. The Effects of Maternity Leave on Children's Birth and Infant Health Outcomes in the United States. J.Health Econ. 30, 221–239. https://doi.org/10.1016/j.jhealeco.2011.01.005

Rowe, J.W., Kahn, R.L., 1997. Successful Aging. Gerontologist 37, 433–440. https://doi.org/10.1093/geront/37.4.433

Sabates-Wheeler, R., Koettl, J., 2010. Social protection for migrants: The challenges of delivery in the context of changing migration flows. International Social Security Review 63, 115–144. https://doi.org/10.1111/j.1468-246X.2010.01372.x

Souchay, C., Isingrini, M., Espagnet, L., 2000. Aging, episodic memory feeling-of-knowing, and frontal functioning. Neuropsychology 14, 299. https://doi.org/10.1037//0894-4105.14.2.299

The World Bank, 2017. Population Growth (No. Population Growth Report), The World Bank, Population Growth Report, United Nations Population Division. World Population Prospects: 2019 Revision. The World Bank, Washington, DC.

Tulving, E., Craik, F.I., 2000. The Oxford Handbook of Memory, 1st ed. Oxford: Oxford University Press, Oxford, England.

UNC Chapel Hill, Carolina Population Center, 2015. Map of Survey Regions (No. Map of Survey Regions), University of North Carolina. University of North Carolina, Chapel Hill, NC.

UNC Chapel Hill, Carolina Population Center, CHNS Research Team, 2010. Survey Design, University of North Carolina. University of North Carolina, Chapel Hill, NC.

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. United Nations, New York, NY.

United Nations, 2013. Population Facts, United Nations, Department of Economic and Social

Affairs, Population Division. United Nations, New York, NY.

United Nations, 2007. World economic and social survey 2007, Development in an ageing world, United Nations, Department of Economic and Social Affairs, E/2007/50/Rev.1. United Nations, Department of Economic and Social Affairs, New York, NY.

Vilela, A., 2013. Pension coverage in China and the expansion of the New Rural Social Pension (No. Briefing no. 11), HelpAge International. HelpAge International, London, UK.

Wang, M., 1993. Living longer is not equal living healthy: Life quality and Medical Security, 1st ed. Beijing: Chinese Economic Press, Beijing, China.

Welsh, K., Butters, N., Hughes, J., Mohs, R., Heyman, A., 1991. Detection of Abnormal Memory Decline in Mild Cases of Alzheimer's Disease Using CERAD Neuropsychological Measures. Archives of Neurology 48, 278–281. https://doi.org/10.1001/archneur.1991.00530150046016

Figures and Tables

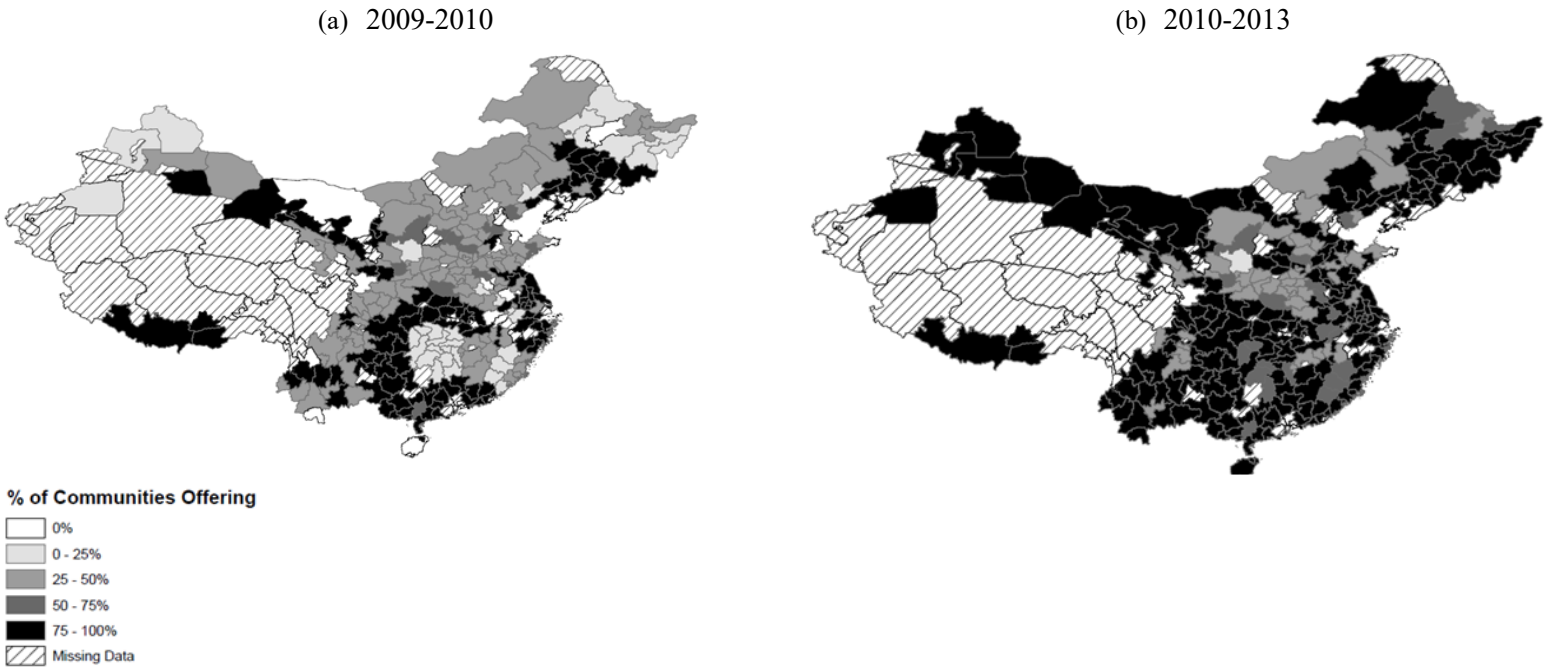


Fig 1. Geographic Implementation of NRPS. This figure shows the timely implementation of NRPS. "% of Communities Offering" indicates the percent of communities (*shequs*) within the province that implemented the NRPS.

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]	Delay Word Recall[a]	Total Recall[a]	Cognitive Memory Index[b]
	(1)	(2)	(3)	(4)
Panel A (ITT):				
Offered NRPS × Above60 [c]	-0.144***	-0.230***	-0.353***	-0.103**
	(0.052)	(0.052)	(0.093)	(0.040)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.230	0.215	0.247	0.313
Observations	21,202	21,202	21,202	21,202
Panel B (TOT):				
NRPS Participation × Above60 [d]	-0.208*	-0.425***	-0.633***	-0.212**
	(0.120)	(0.122)	(0.214)	(0.087)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	241.242	241.242	241.242	241.242
R-squared	0.065	0.152	0.102	0.110
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) 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), 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 3: Actual Retirement and Cognitive Performance.

	Immediate Word Recall[b]	Delay Word Recall[b]	Total Recall[b]	Cognitive Memory Index[c]
	(1)	(2)	(3)	(4)
Retired (Yes=1)[a]	-0.469*	-0.575	-0.906	-0.529**
	(0.283)	(0.369)	(0.613)	(0.252)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
F-Stat (First Stage)	47.48	48.18	47.39	45.82
Beta (First-Stage)	0.135	0.135	0.135	0.135
SE (First-Stage)	0.007	0.007	0.007	0.007
Observations	22,444	22,329	22,226	21,258

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: 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. 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]	Delay Word Recall[b]	Total Recall[b]	Cognitive Memory Index[c]
	(1)	(2)	(3)	(4)
Offer NRPS × Above 60	-0.047	0.375	0.328	-0.052
	(0.487)	(0.482)	(0.874)	(0.378)
Duration of NRPS Benefits (Receiving for 1-2 years)	-0.154	-0.122	-0.276	-0.088
	(0.144)	(0.158)	(0.264)	(0.109)
Duration of NRPS Benefits (Receiving for 3 or more years)	0.204	0.503	0.707	0.333
	(0.316)	(0.364)	(0.534)	(0.225)
Offer NRPS (Yes=1) × Above 60(Yes=1) × Duration of NRPS Benefits (1-2 years)[a]	0.168	-0.046	0.122	0.111
	(0.186)	(0.210)	(0.344)	(0.143)
Offer NRPS (Yes=1) × Above 60(Yes=1) × Duration of NRPS Benefits (3 years or more)[a]	-0.325	-0.727*	-1.052*	-0.382
	(0.339)	(0.393)	(0.582)	(0.246)
Above 60	-0.280	-0.273	-0.554	-0.193
	(0.419)	(0.382)	(0.725)	(0.322)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.234	0.225	0.252	0.310
Observations	22,199	22,092	21,992	21,041

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: 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. 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 Activities				Mental Stimulation Activities		Social Engagement Activities				Health Behaviors And Nutrition			
	#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)	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 per 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 × Above60 [a]	-0.160	-0.024	-1.342**	-1.224**	-0.111***	-0.029**	-0.160***	-0.026	-0.040***	-0.034***	0.157**	-0.019*	-0.020**	17.434
	(0.464)	(0.034)	(0.565)	(0.516)	(0.022)	(0.014)	(0.028)	(0.016)	(0.012)	(0.012)	(0.061)	(0.011)	(0.009)	(39.528)
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	0.186	192.443
Controls	Yes	Yes	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	0.316	0.180	0.356	0.107	0.389	0.257	0.052
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	20,688	23,822
Panel B (TOT):														
NRPS Participation × Above60 [b]	-0.377	-0.055	-3.151**	-2.968**	-0.287***	-0.046**	-0.477***	-0.050	-0.063***	-0.075***	0.354**	-0.042*	-0.045**	37.767
	(1.095)	(0.079)	(1.346)	(1.263)	(0.067)	(0.022)	(0.129)	(0.033)	(0.019)	(0.028)	(0.142)	(0.025)	(0.020)	(85.542)
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	0.186	178.488
Controls	Yes	Yes	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	235.37	234.65	239.65	244.47	223.79	243.02	244.743
R-squared	0.380	0.154	0.357	0.357	0.435	0.263	0.337	0.316	0.178	0.353	0.104	0.389	0.256	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	20,688	23,822

Notes: 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: 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. In constant 2011 Yuan. *p< 0.10, **p< 0.05, ***p< 0.01.

Table 6: 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]	-0.115	0.019	-0.096	-0.058
	(0.095)	(0.121)	(0.192)	(0.070)
Offered NRPS × Above60	-0.033	-0.199**	-0.233*	-0.065
	(0.074)	(0.078)	(0.135)	(0.052)
Offered NRPS × Female	0.029	-0.075	-0.046	0.005
	(0.059)	(0.084)	(0.120)	(0.044)
Above60 × Female	0.029	-0.075	-0.046	0.005
	(0.059)	(0.084)	(0.120)	(0.044)
Above60	-0.481***	-0.429***	-0.910***	-0.302***
	(0.054)	(0.060)	(0.100)	(0.042)
Female	0.076	0.249***	0.325***	0.010
	(0.048)	(0.056)	(0.080)	(0.035)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.229	0.216	0.247	0.314
Observations	21,202	21,202	21,202	21,202
Panel B (TOT):				
NRPS Participation × Above60 × Female[d]	-0.265	0.038	-0.227	-0.134***
	(0.217)	(0.274)	(0.437)	(0.160)
NRPS Participation × Above60[d]	-0.076	-0.452***	-0.528**	-0.147
	(0.168)	(0.176)	(0.305)	(0.118)
NRPS Participation × Female	0.030	-0.074	-0.044	0.006
	(0.059)	(0.084)	(0.120)	(0.044)
Above60 × Female	0.117	-0.092	0.025	-0.075
	(0.100)	(0.125)	(0.196)	(0.077)
Above60	-0.467***	-0.346***	-0.812***	-0.275***
	(0.078)	(0.086)	(0.144)	(0.058)
Female	0.076	0.248***	0.324***	0.009
	(0.048)	(0.056)	(0.089)	(0.035)
Baseline Mean	3.792	2.862	6.678	0.000
Controls	Yes	Yes	Yes	Yes
R-squared	0.228	0.214	0.246	0.313
F-Stat (First Stage)	157.967	157.967	157.967	157.967
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

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 × Above60 × Female [c]	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
	(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 × Above60	-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
	(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 × Female	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
	(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 × Above60 × Female [c]	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
	(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 × Above60	-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
	(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 × Female	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
	(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 Educ	# 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.

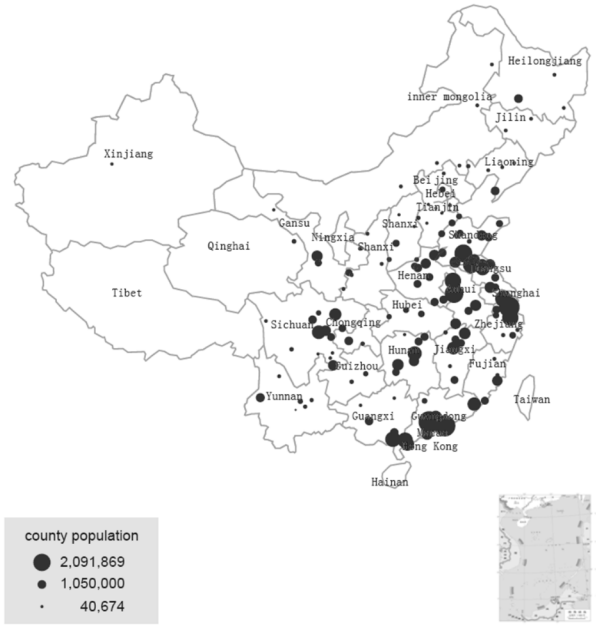
***Significant at the 1 percent level.

**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).

Table B1: 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 B2: 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 B3: 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 ×	-0.228*	-0.472***	-0.700***	-0.221**
Above60[d]	(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 ×	-0.238*	-0.472***	-0.711***	-0.213**
Above60[d]	(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 B4: 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 B5: 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.