ESSAYS ON LABOUR SUPPLY AND HEALTH OF OLDER WORKERS

Yingying Zhang
Department of Economics
Lancaster University Management School

A thesis submitted for the degree of
Doctor of Philosophy

February, 2022

Declaration

I declare that the work presented in this thesis is, to the best of my knowledge and beliefs, original and my own work. The material has not been submitted, either in whole or in part, for a degree at this, or any other university. This thesis does not exceed the maximum permitted word length of 80,000 words including appendices and footnotes, but excluding the bibliography. A rough estimate of the word count is: **51,652**

Yingying Zhang

ESSAYS ON LABOUR SUPPLY AND HEALTH OF OLDER WORKERS

Yingying Zhang
Department of Economics, Lancaster University Management School
A thesis submitted for the degree of *Doctor of Philosophy.* February, 2022

Abstract

The thesis studies three issues: how a change in the expected social pension affects older workers' labour supply decisions under a life-cycle framework; what are the mental and physical health consequences of formal retirement; and how health deterioration affects labour supply transitions of older informal workers. Previous studies on health and labour market transitions focus on older formal workers from developed countries. This thesis adds to the current literature by studying informal workers in China who tend to work until much older ages or until they are physically incapable. They suffer from low levels of income and savings, and limited health insurance coverage for old-age support. Data from the China Health and Retirement Studies are used for all three chapters.

The first chapter, '**Social Pension and Labor Supply Responses of Older Informal Workers**', utilises the community-level difference in the years of introducing the New Rural Social Pension (NRSP) programme in rural China to identify the effect of individual NRSP participation on their probability of working and actual hours of working. We use a trivariate, non-linear random-effects model and adopt life-cycle framework to guide the empirical specification that focuses on heterogeneous responses from participants in different age groups. We find that male participants overall do not change their working probability, while female participants are less likely to work if they are above 50, although the effects are only marginally significant. For participants staying in the labour force, male workers who are below the pension eligible age of 60 significantly increase their weekly working hours by 5.86% to 7.67%, with the effect mainly coming from non-agricultural workers.

The New Rural Social Pension (NRSP) programme is both a basic social security programme that pays monthly basic benefits to age- eligible participants, and a retirement pension programme for age-ineligible participants who are not covered by employees' pension programme. Given that the NRSP is one of the largest social security programme around the world in the terms of enrollees and financial input, evaluating how effective it has been in changing individual behaviours and welfare is important and provides policy implications on developing countries that are reforming or expanding their non-contributory public transfer programmes to cover the larger population of informal workers.

The second chapter, '**Gender Difference in the Retirement Effect on Cognitive Functioning and Depression Risk**', utilises the different compulsory retirement ages for blue collar and white-collar workers to identify both the short-run effect of transitions into retirement and the cumulative effect of retirement years on formal retirees' mental health. We estimate both dynamic and non-dynamic panel data IV models. We focus on the bivariate random effects IV model and find that transition into formal retirement increases men's scores in cognitive tests by 30-50%, but reduces women's scores in the mental intactness test by 52.3%. The positive short-run effects of retirement disappear with the number of years in retirement for men and the negative effect reverses for women. There is no significant retirement effect on either men's or women's depressive risk or physical well-being.

The third chapter, '**The Dynamic Relationship between Depression, Physical Health, Labour Market Exits and Entries of Older Informal Workers**' studies the sequential causality between mental and physical health and labour market transitions of older informal workers, while also identifies the roles of state dependence, observed SES variables, and unobserved individual heterogeneity in explaining health and labour supply outcomes. Using a trivariate dynamic cross-effects model, we find that depression increases the probability of male nonworkers re-entering the labour force, mainly by taking agricultural works, in the subsequent period. Physical health deterioration predicts a higher risk of depression in the subsequent period, larger for men than for women. Depression does not have any effect on men's reported number of physical problems in the subsequent period though increases women's, supporting an asymmetric effect between physical health and mental health. Physical health declines predict an exit from the labour force, with sectoral and gender-specific difference. There is strong and significant state-dependence for the number of physical health problems and labour market status, but not for the risk of depression, similarly for men and women. The paper provides empirical evidence that adding mental health into studying the relationship between health and labour supply transitions can better explain the variation in health effects and the mechanisms behind the labour supply adjustments after health shocks.

Acknowledgements

First and foremost, I would like to thank my supervisors, Steve Bradley and Robert Crouchley, for their dedicated supervision, critical and valuable feedback, professional and technical guidance. Besides working together on my PhD projects, Rob has also been an inspiring a good teacher when we first met in my master program, and a good colleague when I worked as a teaching assistant for him. Both Steve and Rob have been very supportive since I entered the job market.

I would also like to thank Maria Navarro Paniagua for being a great internal examiner. Her helpful comments and kind words have encouraged my a lot. I also feel grateful for the supports I have received from the Department of Economics at Lancaster University throughout the past years. I would like to especially thank Marwan Izzeldin for his support and guidance in my teaching activities and Caren Wareing and Sarah Ross for their administrative work.

Besides the Department of Economics, I want to thank the Mathematics and Statistics Hub (MASH) Team for providing a lovely and warming working environment. I enjoy working as a tutor there. I want to especially thank Anna Karapiperi and Eleftherios Kastis for their support and empathy during the past difficult years. I cannot express my full gratitude for the Lancaster University Mangement School (LUMS). Without the LUMS scholarship and institutional support, this study would not have been completed.

Finally, I want to thank my friends and housemates for their emotional supports, thank my fellow PhD students for their companions, and valuable feedback throughout my PhD study. Among them, I particularly want to thank Hankui Peng, Huan Yang, Afnan Al-Malk, Gerda Buchmueller, Alex Skouralis, Ugur Satic, Mbakile Seabe, Alex Farnell for their friendship and emotional support throughout the past few years. I would also like to thank Manjeet, Vihba, Lili and Sam for their companionship throughout the difficult and isolated lock-down periods. Above all, my sincere gratitutes go to my family for their unconditional love and full support. I could not imagine how this study would have been completed without their encouragement. I would like to dedicate the thesis to my parents.

Contents

List of Figures

List of Tables

Chapter 1

Social Pension and Labour Supply Responses: Evidence from New Rural Social Pension in China

1.1 Introduction

In the 21st century, China has been reforming its welfare system by introducing a range of universal basic social security programmes, including monetary and in-kind transfers, health insurance and old-age pension that aim to alleviate and prevent extreme poverty, focusing on rural areas and rural and informal workers who were not covered by any public medical scheme or retirement pension scheme. Among these programmes, the New Rural Social Pension (NRSP) programme is special and important in the sense that it is both a basic social security programme and a retirement pension programme for informal or temporary workers who concentrate in the agricultural sector, and non-public, small and medium-sized enterprises in the non-agricultural sector. These informal workers do not have employers to contribute to their pension scheme and lack of financial literacy to plan for retirement themselves, and are thus more vulnerable to financial or health risks when they get older. In China, older rural or informal workers rely mainly on savings and private transfers from adult children for old-age support. The New Rural Social Pension programme was introduced in 2009 and reached nation-wide coverage by 2012. It is one of the largest social security programme in the world, and the 2011 State Council report shows that 326 million rural residents participated in the NRPS and among them, 85 million received pension benefits.

Utilising the community-level difference in the years of introducing the NRSP, the paper studies the pension participation on labour supply decisions of older informal workers, heterogeneous for people in different life-cycle stages and separately for men and women. We also examine the contribution decision of pension-ineligible participants and relate it to the participation decision and labour supply decision. The answers to these questions are important for several reasons. Firstly, examining pensioners' labour supply responses provide suggestive evidence for a lifecycle consumption smoothing behaviour of older, informal workers after an external shock to expected social security wealth. Participants aged above 60 when the programme starts in their residential places are eligible to receive a monthly payment of basic benefits without making any compensatory contribution to the scheme. For them, the pension programme generates a pure income effect and plays a similar role as non-contributory public transfer programmes do. Before China, developing countries such as Brazil (de Carvalho Filho, 2008), South Africa (Ardington, Case and Hosegood, 2009; Duflo, 2000; Jensen, 2004), Mexico and India (Kaushal, 2014) have also introduced or expanded social pension programmes targeting on the elderly. Nonetheless, these programmes are means-tested and offer much higher benefits than the NRSP does. The basic benefit levels of the NRSP remain low and account for only about 10.0% of the average personal disposable income of rural residents. Therefore, evidence from these countries cannot apply directly to the case of the NRSP. Evidence on the NRSP provide policy implications on developing countries that are reforming or expanding their non-contributory public transfer programmes to cover the larger population of informal workers.

Secondly, given that the NRSP is one of the largest social security programme around the world in the terms of enrollees and financial input, evaluating how effective it has been in changing individual behaviours and welfare, and how well it has encouraged savings and contributions into pension accounts, are of great importance in itself. The paper contributes to existing studies of the NRSP by developing a novel identification strategy and estimating the three processes of pension participation, labour market participation, and hours of work together in a trivariate error-correlated model. We study not only the average treatment effect of receiving pensions on the labour supply behaviours of pensioners, but also the effect of contributing to individual accounts in an expectation of future pensions receipt on the current labour supply behaviours of younger participants. About 75% of the age-ineligible participants contribute at the minimum required level to their pension accounts[1]. We expect them to behave differently in the labour market compared with people contributing at a higher level, due to liquidity constraints, time preference and expectation about old-age support. The longer span of data we use captures some changes in contribution levels over time and thus enables us to study the labour supply adjustment in accompany with a change in contribution levels, especially a change from the minimum

[1]The low incentive of younger workers to participate into and contribute in a higher level to the pension scheme can be explained by the low rate of return on the individual investment it offers, low trust in the government, individual liquidity constraints and time preference (Lei et al., 2014*b*).

contribution to a higher level. Studying the age-ineligible group provides long-run implications of the NRSP on retirement preparation of the targeted group.

Another contribution of the paper is adopting a life-cycle framework to guide the empirical specification that focuses on heterogeneous responses from participants in different age groups. Within the framework, lifelong consumption is financed by labour earnings during the working period, by social pensions, private savings and private transfer during the non-working period. Given that leisure is a good and that rural or informal workers have discretion over working hours and time to stop working under a temporary, flexible contract or in many cases no contract, a positive shock to their pension incomes is expected to substitute for labour income and see a decline in hours of working or exit from the labour force. Nonetheless, the substitution effect can also work on private savings and encourage more consumption, or reduce private transfer from adult children, which is not found to be significant in Huang and Zhang (2021)'s study. Therefore, the labour supply behaviour of the pensioners is also an empirical issue. Behavioural change of age-ineligible participants is also an empirical issue and people approaching 60 is expected to respond more than younger participants due to lower uncertainty of future pensions and survival risks, liquidity constraints and time preference. Accrual effect means that they may work more to save more on their pension accounts. Empirically, we study age-specific labour supply responses to the participation in the NRSP by interacting individual pension participation status with an exhaustive set of age group dummies.

Last but not least, the paper adds to the empirical evidence on evaluating and reforming pension programmes and non-contributory public transfer programmes. In contrast to the positive shock to social security wealth generated by the introduction or expansion of social pension programmes in developing countries, reforms on state pension and retirement pension programmes in developed countries facing the ageing population have created negative wealth shocks to the affected employees. All these shocks have been found to affect individual labour supply behaviours and private savings to some extent (Attanasio and Rohwedder, 2003; Atalay and Barrett, 2015; Hanel and Riphahn, 2012), developing countries:(de Carvalho Filho, 2008; Bertrand, Mullainathan and Miller, 2003; Jensen, 2004; Posel, Fairburn and Lund, 2006; Ardington, Case and Hosegood, 2009). Older workers in low socio-economic status are found to be more sensitive to changes in their expected social security wealth due to their greater dependency on pension and health insurance.

We find that male participants overall do not change their working probability, while female participants are less likely to work across all age groups. For participants who stay in the labour force, male workers below 60 increase their weekly working hours by 5.86% to 7.67%. The effect mainly comes from non-agricultural workers, who overall contribute more to their pension accounts. In contrast, female participants aged above 50 start to reduce their hours of working. Cross-equation covariances of the error terms suggest that pension participation is endogenous to women's probability of working as there are unobserved effects that predict both participation in the NRSP and staying in the labour force for women. For men, pension participation is endogenous to their hours of working. Heterogeneous analyses are conducted on individuals from different job sectors, educational groups, and asset quantiles.

The structure of the paper is as follows. Section 2 introduces the institutional background of the China pension system and the New Rural Social Pension programme. Section 3 reviews on the literature. Section 4 discusses the life-cycle framework that guides the empirical specification and interpretation. Section 5 introduces the data and describes demographic statistics. Section 6 specifies the econometric models, and Sections 7 and 8 report the main results. We report in Section 9 some robustness check results and conclude the paper in Section 10.

1.2 Pension System in China

1.2.1 History of Pension Provision in China

Before the introduction of the NRSP in 2009, the old pension system in China was comprised of the largest monolithic urban basic occupational pension scheme (pillar I) that leans heavily towards workers in government or state-owned companies in terms of level of pension benefits, the much smaller-coverage, voluntary occupational and personal savings arrangements (pillar II) such as the Enterprise Annuity, and under-developed, commercial pension products (pillar III)[2]. According to the State Council report in 2017, the urban basic occupational pension scheme has 379 million participants, accounting for 91.55% of the total urban employees. Among these participants, only 23.25 million participate in enterprise annuity schemes. Rural non-waged workers and rural migrant workers who generally work under temporary or no labour contracts are uncovered by any retirement pension scheme. These uncovered populations used to rely on a high level of precautionary savings and private transfers for old age support. Nonetheless, the rapid demographic changes and aging population as a result of declining fertility and rapid urbanization have challenged the traditional ways of old-age supports, especially in rural China, justifying the introduction of a widely-covered, heavily government-subsidized, basic social pension scheme for rural or informal workers.

The NRPS was introduced in 2009, and the Urban Residents Pension Scheme

[2] Before 2015, the urban basic occupational pension scheme was comprised of the Basic Old Age Insurance covering employees in registered enterprises (public or private), and the Public Employee Pension scheme that offers pensions at a much higher replacement rate to civil servants and workers in non-profit government institutions such as schools and hospitals. The two were merged into a uniform programme for all urban workers (Fang and Feng, 2018)

(URPS), a paralleled programme for urban workers not covered by the employees' occupational pension, was launched later in mid-2011[3]. The two schemes reached nation-wide coverage by the end of 2012 and were unified officially into the Universal Residents Social Pension Scheme (RSP) in 2014. A more detailed introduction to the modern pension system in China is available in Fang and Feng (2018).

1.2.2 Recent Developments: the NRSP

As shown in Table 1.1, statistics from the Ministry of Human Resources and Social Security (MOHRSS) of China, the coverage of the RSP is larger than that of the long-existing urban employees' occupational pension, and more than 400 million out of the 512.55 million participants in the RSP are rural workers, or participants in the NRSP. Among the 401.99 million of urban employees covered by the employees' occupational pension scheme, about 37 million work in the public sector. However, the RSP offers an average benefit lower than 10% of the average benefit of the urban employees' occupational pension programme, primarily due to the relatively low contribution levels of the rural or informal workers. Based on the 2018 China Statistics Yearbook published by of Statistics of China (2018), the average annual household per capita disposable income amounts to 13,432 RMB for rural residents, about 37% of the 36,396 RMB for urban residents. Therefore, more efforts should be made in incentivizing and educating rural workers to save more in their pension accounts.

Table 1.1: Comparison of Employees' Pension and Residents' Social Pension

	Urban Employees' Pension	Residents' Social Pension
Participants (million)	401.99	512.55
Pension Fund (billion RMB)	4279.39	328.8
Pension Payment (billion RMB)	3792.43	239.87

Notes: Based on 2017 Annual Report of the Ministry of Human Resources and Social Security of PRC

In terms of pension policy, the RSP differs from the urban employees' pension programme in that it is disconnected from employment and subject to voluntary participation. Eligible individuals, or workers not covered by the urban employee pension scheme, opt in and contribute a voluntary amount of money to the pension scheme, subject to a minimum requirement level. The urban employees' pension programme requires employers to contribute 20% of employees' wages to the social pension pool that funds the basic benefits and employees to contribute 8% of their wages to individual pension accounts. The eligible age to retire and receive pension

[3]Pilot programmes started in 2008 in few areas including counties in Suzhou (Jiangsu Province), suburban areas of Shanghai, and Baoji (Shaanxi Province) (Zhang and Tang, 2008).

is 50 for female blue-collar workers, 55 for female white-collar workers and male blue-collar workers, and 60 for male white-collar workers. In the RSP, central and local governments take the role of employers and pay for basic benefits, though at a level far below the poverty line. Individual contribution is tiered and not tied to their wages. The eligible age to receive a pension is 60 for both males and females, and is not related to the termination of a labour contract like it is in the urban employees' occupational pension scheme. These special rules make the RSP both a retirement pension programme and a social security programme.

Pension benefits distributed after at least 15 years of contribution are comprised of the basic benefits subsidized by the central (55 RMB per month, or 8.1 dollars based on the 2011 exchange rate) and local governments (varying based on local economic status), and the accumulative contributions divided by an annuity factor of 139. The minimum required amount of contribution is 100 RMB per year in most provinces. To encourage participation and higher levels of contribution, local governments provide subsidies that match and slightly increase with the levels of individual contributions.

Participants aged above 60 when the pension scheme started can receive basic benefits without making compensatory contributions to the scheme. According to an official statement published by The Ministry of Human Resources and Social Security, the basic benefit averaged across 23 provinces (not including the four municipalities of Beijing, Shanghai, Tianjin, Chongqing where basic benefits are substantially higher than the rest of the country) is about 90 RMB per month in 2015, or 0.48 dollar per day, which represents 25.3% of the international poverty line of 1.9 dollar per day, or 1,080 RMB per year, 38.6% of the national poverty line of 2,800 RMB per year. Based on the average per capita annual disposable income for 2018 (of Statistics of China, 2018), which is 10,772 RMB, the average basic benefit of the NRSP accounts for only about 10.0% of the average personal disposable income of rural residents.

For the age-ineligible NRSP participants in our sample who are required to contribute until they reach the pension-eligible age of 60, 75% of them contribute at the minimum level of 100 RMB per year, which is matched with a 30 RMB subsidy according to policy rules. Assume a typical contributor pays for his personal account for 15 years. The accumulative contributions will be paid out based on an annual factor of 139 after the age of 60, so the pension benefits he receives would equal to 90+(130*15)/139=104 RMB per month, or 1,248 RMB per year, 11.6% of the average rural disposable income in 2015. For those remaining in the labour force after 60 and receiving labour income, the amount of pension benefits accounts for an even lower percentage. We do not consider the factors of time preference, uncertainty and survival risk here.

Table 1.2 shows the decomposition of rural household disposable income in 2017, based on the China Yearbook of Household Survey. It provides evidence that the pension incomes contribute to, on average, only about 5% of the total disposable

Table 1.2: Decomposition of Rural Household Disposable Income (2017)

	2013	2014	2015	2016
Wages	38.7	39.6	40.3	40.6
Farm-work income	30.1	28.6	27.6	26.4
Non-farm-work income	11.6	11.8	11.8	11.9
Capital income	2.1	2.1	2.2	2.2
Pension income	4.3	4.7	5.3	5.7
Other transfer income	16	16.1	16.1	16.5
Pension payment	-2.1	-2.2	-2.7	-2.8
Other transfer payment	-0.7	-0.7	-0.6	-0.6
Disposable income (%)	100	100	100	100

Notes: All numbers are percentage points. Data come from the 2017 China Yearbook of Household Survey.

income of a rural household, or 12.5% of wages, or 15 to 20% of farm work income.

Overall, given the low income replacement ratio of the NRSP, we expect overall negligible labour supply responses from the enrolled workers. The pension income may have a bigger impact on the low-income group, such as the very old agricultural workers, and the low-educated group. Apart from the pure income effect, the pension effect we estimate also picks up the effects of the pension programme on motivating savings and on financial education, especially for middle-age workers.

1.3 Theoretical Framework

Following Attanasio and Rohwedder (2003) and Feng, He and Sato (2011), we use the life-cycle model as a conceptual framework to guide the specification of our econometric models and the interpretation of estimation outputs. Within the framework, an individual adjusts labour earnings and pension wealth to finance his/her life-long consumption.

1.3.1 Pension Wealth

For a single individual aged below 60, the expected value of pension wealth at time t, $E_t(P_t)$, is defined as the present value of pension benefits minus the present value of current and future pension contributions, inspired by Attanasio and Rohwedder (2003):

$$E_t(P_t) = \sum_{k=60}^{T} \frac{b_k * s_k}{(1+r)^{k-t}} - \sum_{k=t}^{59} \frac{con_k * s_k}{(1+r)^{k-t}}, t < 60 \tag{1.1}$$

Where b_k denotes the expected annual pension benefits individuals receive in time k after the pension-eligible age of 60, s_k is the probability of surviving till year k. T is the maximum attainable age. con_k is the annual contribution before 60, r is the one-year real interest rate.

As an individual aged 60 or above when the NRSP starts can receive the basic pension without making any contribution, his expected pension wealth simplifies to:

$$E_t(P_t) = \sum_{k=t}^{T} \frac{B_k * s_k}{(1+r)^{k-t}}, t \geqslant 60 \tag{1.2}$$

Where B_k is the annual basic benefits in year k. Basic pension benefits subsidized by local governments are heterogeneous across regions and are expected to increase over time with economic development.

1.3.2 A Simple Model

The simplified life-cycle model does not explicitly account for uncertainty, bequest motive, and dynamic discounting rates, as our data and settings do not enable us to separately identify these effects. In our simple model, pension wealth as an alternative source of financing lifetime consumption, has an offset effect on labour earnings, private savings, and private transfers. We focus on the offset effect of pension wealth on labour income. Expected pension wealth should be viewed as equal to future disposable earnings, though wage earnings are more liquid, and pension benefits accrue at a more stable rate (Gruber and Wise, 1998). The offset effect of pension wealth on other wealth has been examined in previous studies (Gale, 1998; Attanasio and Rohwedder, 2003). Specific to the case of China, Feng, He and Sato (2011) studied the urban employees' pension reform from 1995 to 1997 that substantially reduced the expected SSW for contributors, and found the reform boosted household savings rates by 6 to 9% for cohorts aged 25–29 and by 2 to 3% for cohorts aged 50–59.

Following (Feng, He and Sato, 2011), suppose an individual at time period t chooses current and future consumption to maximize lifetime utility. Assume an iso-elastic utility function, the maximization problem becomes:

$$\max \sum_{k=t}^{T} \beta^{k-t} logC_k \tag{1.3}$$

The lifetime budget constraint includes discounted current and future labour incomes, private transfers, pension benefits, and current household assets:

$$s.t. \sum_{k=t}^{T} \frac{C_k}{(1+r)^{k-t}} = \sum_{k=t}^{TR-1} \frac{E_k(1-\tau_t)}{(1+r)^{k-t}} + \sum_{k=TR}^{T} \frac{TF_k}{(1+r)^{k-t}} + E_t(P_t) + A_t \tag{1.4}$$

where k indexes age or time, C_k is consumption in each period, β is the time preference rate, r is the real interest rate. E_k is the real labour earnings in each period. τ_k is the proportion of earnings transferred to children or elder parents when in the labour force, which is also specified in (Zhao, Li and Chen, 2016)'s model. $E_t(P_t)$ is the discounted future pension benefits received upon 60, as specified in equations (1) and (2). TF_k is the net private transfers from the family, mainly from adult children, after exit from the labour force. A_t is the household assets at time t, including savings and capital income in previous periods and values of capital assets such as houses, land, agricultural products, and the household durable assets. T is the maximum attainable age and TR is the age he or she stops working.

Current consumption and labour income decisions can be specified as:

$$C_t = \frac{1-\beta}{(1-\beta)^{T-t+1}}[\sum_{k=t}^{TR-1} \frac{E_k(1-\tau_t)}{(1+r)^{k-t}} + \sum_{k=TR}^{T} \frac{TF_k}{(1+r)^{k-t}} + E_t(P_t) + A_t] \tag{1.5}$$

$$C_k = (\beta(1+r))^{k-t} C_t \tag{1.6}$$

Assumptions: 1) workers stop working completely after the age of TR; 2) pension participants below 60 believe themselves to outlive the pension-eligible age of 60.

Equation (5) shows that in the absence of the pension programme, the consumption of rural informal workers is mainly financed by labour income when in the labour force and by savings during working life and private transfers after exit from the labour force. After the introduction of the NRSP scheme, pension participants expect an extra source of income, pension benefits, after the age of 60, which does not necessarily overlap with their age of stopping to work. Equation (5) also suggests that an increase in future SSW will crowd-out private savings during working life, encourage current consumption, and may probably affect labour supply decisions given the flexibility of informal jobs. As a result, workers can increase current consumption or reduce labour supply or do both, depending on the marginal utility gained from extra consumption relative to the utility of leisure. Empirically, we would expect the labour supply responses to vary for individuals in different stages of their life-cycles when the pension programme is introduced, as suggested by Attanasio and Rohwedder (2003) and Feng, He and Sato (2011), in different income groups and working in different job sectors.

Elder participants may respond more because they: 1) have shorter discount periods and higher present value of basic benefits given the same entitlements as younger participants; 2) have lower uncertainty about survival risk; 3) contribute less than younger cohort participants do. Previous empirical studies using a life-cycle framework to estimate behavioural responses to pension reforms (Attanasio and Rohwedder, 2003; Feng, He and Sato, 2011) tend to find smaller effects on the younger cohorts than on the older cohorts. Nonetheless, the older cohorts also have smaller SSW because they: 1) have fewer years to build up entitlements; 2) have lower earnings and probably make lower levels of annual contribution than younger cohorts do. Comparison of the sizes of responses of participants in different age groups is an empirical issue.

A stable stream of earnings history is not available due to the short span of the survey data we use. Farmers and informal workers, the target group of the NRSP, tend to have flexible working hours and unstable labour earnings. As an alternative to labour earnings, we estimate working hours and assume stable productivity or hourly wages. The fact that the average basic benefit is about 10.0% of rural personal disposable income provides a guideline to interpret the magnitudes of labour supply responses in empirical estimation.

In the empirical specification, we estimate individual labour outcomes as functions of pension participation status or annual contribution level, instead of accumulated expected pension benefits. One reason is that pension participation is equivalent to receiving basic benefits, which is universal across the countries with little regional difference, for people above 60. More than 75% of the younger participants contribute at the minimum requirement level to their pension accounts, and they would expect to receive similar level of pension incomes after 60. To reveal the life-cycle pattern of labour responses, we interact the NRSP participation status with an exhaustive set of age-group dummies to study cohort-specific responses. We control for observable individual-, household- and community-level characteristics, and account for unobservable factors that affect both individual labour supply decisions and pension participation decisions. Estimations are separated for males and females due to different gender roles in rural China, and their systemically different attachment to the labour market. A significant labour supply response provides evidence that a growth in the social security wealth to some extent crowds out labour incomes.

1.4 Related Literature

Extensive theoretical and empirical studies from developed countries find that institutional structures of benefits and pension programmes are prime determinants of retirement (Gustman and Steinmeier, 1983, 2005; Rust and Phelan, 1997; Heyma, 2004; French and Jones, 2011). Theoretically, the wealth effect enables individuals to finance retirement with fewer years of labour supply. The illiquid nature of pension benefits, available only upon the eligible age and cannot be borrowed against with, leads to a high job exit rate in early retirement age against normal retirement age (Rust and Phelan, 1997; Gustman and Steinmeier, 2005). The accrual effect encourages labour supply because an extra year of working increases contributory earnings and expected social security benefits. The magnitudes and relative importance of these effects on personal retirement decisions depend on the specific setup of a pension programme, individual and job-related characteristics.

A lot of factors have been introduced in structural retirement and pension models to explain the observed labour supply responses, including survival and health uncertainty, wage uncertainty (French, 2005); capital and health insurance market imperfections (Rust and Phelan, 1997; French, 2005); spill-over effects from spouses' labour supply behaviour and shared leisure (Gustman and Steinmeier, 1994), spouses' health and skills (Van der Klaauw and Wolpin, 2008). These factors may have different roles in explaining the labour supply responses of informal workers in developing countries due to the different institutional backgrounds and pension policy rules. The lack of evidence using developing countries' data, and the complexity of the theoretical relationship between changes in pension wealth and behavioural responses make it an empirical issue to evaluate the labour supply responses to the NRSP. Another difference between informal or rural workers in China and formal workers in developed countries or urban China is that they have more discretion over hours of work and whether to work, though they are subject to little job security. We focus on non-working and not permanent exit from the labour market or retirement in this paper, as it is common for informal workers to exit and re-enter the labour market when faced with health or financial shocks.

Regressing labour supply on social security wealth can generate bias as social security wealth is a nonlinear function of average lifetime earnings and correlates with unobservables that determining retirement. Studies exploit natural experiments such as exogenous pension policy changes and again focus on developed countries (Coile, 2015). Using a DiD approach, Atalay and Barrett (2015) study 1993 Australian Age Pension reform that increases the pension-eligibility age for Australian women. They find the retirement probability declines by 8 to 13% while the probability of enrolling in other social insurance programmes increases. Given that the Australian Age Pension programme is a non-contributory scheme not conditional on employment or earnings history, it provides evidence of a pure wealth effect on labour supply. In the case of the NRSP, the wealth effect is also expected to be the main driving force behind pensioners' labour market transitions

Mastrobuoni (2009) studying the US data finds that the mean retirement age increases by about half as much as the increase in the Normal Retirement Age. Hanel and Riphahn (2012) study a similar reform that raises the normal retirement age for women in Switzerland find that a 3.4% decline in retirement benefits reduces age-specific probability of retirement by 50%. In summary, empirical evidence using data from developed countries suggests that retirement pensions have significant and substantial effects on workers' retirement behaviours, with the effect and the elasticities being larger for older workers than for younger workers (French and Jones, 2011).

Our study is more relevant to studies of social pension programmes in developing countries such as Brazil, South Africa, Mexico due to their similarly large population of informal workers and institutional background. Low labour income levels and low savings, imperfect credit markets, and limited coverage of the social safety net mean that informal workers in developing countries are vulnerable to health and economic shocks and tend to work until older age.

de Carvalho Filho (2008) studies behavioural responses to the 1991 Brazilian social security reform that lowers the pension eligibility age from 65 to 60 and doubles benefits for rural workers. He finds that the affected male workers receive benefits, are 38% less likely to work, and reduce weekly working hours by 22.5 hours. The Brazilian rural pension is similar to the NRSP as both are non-means-tested programmes targeted at rural workers. Pension benefits of both programmes are not conditional on retirement or income history, but on individual contributions and local governments' subsidies. Nonetheless, the Brazilian rural pension is much more generous and amounts to 100% of the minimum wage after the reform, while the NRSP offers a basic benefit of only 38.6% of the national poverty line. We expect to see different levels of labour supply adjustments to the NRSP.

The South African Old-Age Pension (SAPP) Programme is a non-contributory but means-tested government transfer programme that pays about twice the average per capita household income to eligible women over 60 and men above 65. Related studies have focused on intra-family allocation of pension benefits and the resulting labour supply behaviours of household members, especially of prime-age members. Jensen (2004) finds no significant change in home or migrant labour supply, but a one rand increase in pension income is met with a 0.25–0.30 rand reduction in private transfers. Bertrand, Mullainathan and Miller (2003) find that a 1000 rand change in pension income reduce the working probability by 9.9% and working hours by 17 for 16-50 year old or prime-age household members of pensioners. Posel, Fairburn and Lund (2006) also find a 6% increase in the probability of females being migrant workers when there are female pensioners in the households. Ardington, Case and Hosegood (2009) also find that receiving pension increases children's probability of being migrant workers by 7%. For pensioners themselves, Ranchhod (2006) uses the regression discontinuity method and finds a reduction in labour supply by 8.4% for men and 12.6% for women. Given that the NRSP offers much less generous benefits than the SAPP does, we do not study the spill-over effects on other household members and focus on pensioners' behavioural changes. Nonetheless, there exist studies looking at these outcomes of the NRSP and they tend to find no or negligible spill-over effects on other household members (Huang and Zhang, 2021).

Previous studies of the NRSP focus on the short-run, pure income effect of receiving pensions on behavioural changes of pensioners, covering outcomes of incomes and food expenditures, retirement, health and health behaviour, private transfers and extended families' behaviours. Some studies use cross-sectional data and the county-by-county roll-out of the NRPS to estimate intention-to-treat effects on the age-eligible rural population (Li, Wang and Zhao, 2018; Huang and Zhang, 2021), while others use panel data and community-level duration of the NRSP as an instrumental variable for individual participation to study average treatment effects (Cheng et al., 2018*a*,*b*; Li, Wang and Zhao, 2018; Shu, 2018). These studies tend to find that receiving pensions reduces both the probability of staying in the labour force and hours of working for the age-eligible or the pensioners. The effects are more pronounced among agricultural workers and low-income groups, though inconclusive for different gender groups. Very few studies look at the behavioural responses of younger, age-ineligible workers (Shu, 2018; Huang and Zhang, 2021; Chen, Zhang and Zhang, 2017) and they tend to find a null pattern among the group. Only one study (Chen, Hu and Sindelar, 2020) investigates determinants of contribution levels.

Specifically, utilising the county-by-county roll-out of the NRPS, Huang and Zhang (2021) estimate intention-to-treat (ITT) effects of the NRSP on various outcomes for both the age-eligible and the age-ineligible. They find that receiving pension increases household income by 17.6% and food expenditures by 9.6%. Participation in the NRSP reduces agricultural labour supply by 3.6% for pensioners and 5.8% for younger participants, who are 3.3% more likely to take up non-agricultural jobs. They also find an improvement in health status for pensioners in terms of a 3.2% decline in disability and 1.7% in malnutrition, and a 2.17% lower mortality rate. The paper differs from their study and estimates average treatment effects (ATE) of the NRSP on individual labour supply behaviour, not only on working probability but also on hours of work. We separate the analysis for gender due to their systemically different level of attachment to labour market, and study the NRSP effect by age groups to implicate on a life-cycle pattern of the offset effect of pension wealth on labour income. We utilize community-by-community roll-out of the NRPS to predict individual participation in the NRSP.

Some studies using the same data from CHARLS to evaluate the NRSP effect on labour supply tend to find that both pensioners and contributors retire earlier than they otherwise would do, though the evidence is diverse in terms of magnitude due to different empirical strategies (estimating ITT versus ATE) and specifications adopted. Nonetheless, none of the studies look at age-specific heterogeneity in labour supply responses, which is important in interpreting life-cycle pattern of individual consumption smoothing behaviour. For the very few studies estimating the ATE of the NRSP on individual labour supply, Shu (2018) does not account for the truncated nature of working hours variable and the estimated results would suffer from selection bias. Bias is also likely to come from the fact that he includes endogenous health variables to explain retirement behaviour. The identification method is also different and we account for the binary nature of working and pension participation.

Using a fixed-effect difference-in-difference model, Li, Wang and Zhao (2018) estimate the ITT of the programme and find that rural, age-eligible men reduce agricultural working hours by 13.5% and substitute for the less physically demanding task of caring for grandchildren. Taking a RE IV approach and using community-level duration of the NRSP as an instrument for individual decision to participate in the NRSP, Shu (2018) estimates the ATE and finds that receiving pensions significantly reduces the probability of working by 27%, and reduces working hours and agricultural hours per week by 56.4% and 69.7% respectively for males and females. Male and female participants yet to receive pensions are 14.2% and 17.3% less likely to work. Using the same identification strategy but different dataset, Cheng et al. (2018*a*) find that pensioners are 6.2% more likely to retire and substitute for household chores, grandchild care, social and leisure activities. In summary, studies of the NRSP suggest that receiving pensions significantly increases the retirement probability and reduces hours of working, especially for agricultural workers. Only two studies look at age-ineligible participants or potential participants (Shu, 2018; Huang and Zhang, 2021), and they have different conclusions on the labour supply behaviour of this younger group.

The paper differs from the studies mentioned above in that we focus on labour supply transitions of participants, and do not assume that they retire and never come back to work. We use an alternative categorization of being economically inactive or non-working instead of retirement, to account for the fact that substantial amounts of respondents who reported being retired in earlier waves going back to the labour force in later waves, especially for those leaving from agricultural jobs. We also implicate on the life-cycle responses by estimating the effects on participants from different age groups. Last but not least, we contribute to the literature by using a novel methodology to estimate together the processes of participation in the NRSP, working probability and working hours in a trivariate error-correlated model. The model can better account for the confounders that affect both labour supply decision and pension participation decision than other more conventional approaches adopted in the existing studies of the NRSP.

1.5 Empirical Specification

We can compare the labour supply outcomes of participants with those of non-participants, given that both groups come from communities that have started the NRSP. Alternatively, we can compare participants with non-exposed individuals coming from communities that have not yet started the NRSP. Previous studies using duration of the programme as an IV do not differentiate between non-participants and the non-exposed (Cheng et al., 2018*a*; Shu, 2018). The duration of programme my only capture the year effect that would also affect individual labour supply behaviours. Therefore, we use an indicator of the availability of the NRSP as an instrument. By doing so, we are comparing labour supply behaviours of NRSP participants in villages that have introduced the NRSP and non-participants in villages that have not introduced the NRSP, given that the two groups share similar observed and unobserved characteristics. According to Wooldridge (2010), the endogenous variable and the instrumental variable can be binary variables. To account for the binary nature of the endogenous variable of NRSP participation, and the binary nature of the working probability, we conduct simultaneous equations analysis and estimate probit models for the binary dependent variables.

As participation in the pension programme is voluntary, self-selection can be a concern, and older people approaching the pension-eligible age are more likely to participate in the the NRSP than the younger and poorer workers are. Two strategies have been used in the literature to tackle the bias. The first approach focuses on pension eligibility and estimates the intention-to-treat (ITT) effect instead of the ATE of actual participation in the programme (Bertrand, Mullainathan and Miller, 2003; Li, Wang and Zhao, 2018; Huang and Zhang, 2021). The NRSP was introduced in different communities in different years, and the availability of the NRSP in local communities/ counties has been utilized to study the ITT effect on individual or household outcomes. The assumption is that eligible people participate in the pension programme once it is made available in local communities, or are affected by the programme.

The second approach finds an instrument for individual participation that is exogenous to other individual outcomes. Previous studies have used the county-level, heterogeneous duration of the NRSP in local counties/ communities to estimate an ATE effect of participation in the NRSP (Cheng et al., 2018*a*; Shu, 2018). The assumption is that the longer the NRSP has been in place, the better knowledge people would have about it and the more commitment they would build on it, therefore the more likely they are to participate. Another potential instrument is the distance to the pension-eligible age of 60. Previous studies using fuzzy regression discontinuity (RD) have used the pension-eligible age of 60 as a cut-off point to study the effect of receiving the new rural pensions on various individual and household outcomes. The approach is similar to using the pension-eligible age of 60 and individual age normalized to be zero around 60 as instruments for the probability of receiving pensions. The same set of instruments can be used for predicting participation in the NRSP, under the assumption that individuals approaching or above the pension-eligible age are more likely to participate than those further below 60.

The CHARLS survey only releases information about the location of sampled individuals at the prefecture level, a level higher than the county level. We therefore identify the NRSP placement time based on the 2011 community survey and official documents published by the Ministry of Human Resources and Social Security. Details of how we construct the instrument are available upon request. The main discussion of the paper is based on the estimation of the ATE. Specifically, we compare labour supply behaviours of participants in communities that have started the NRSP (the treated) and the non-participants in communities that have not started the NRSP (the non-exposed), using the community-level placement of the NRSP as an instrument for individual pension participation. In Appendix A, we also estimate the ITT effect by regressing labour supply outcomes on the availability of the NRSP in local rural communities.

Endogeneity can also come from sample selection when we study individual adjustments to their hours of working, as we only observe the working hours of current workers. People planning to stop working in the near future may gradually reduce hours of working, and the older part-time workers who work fewer hours than full-time workers do are more likely to reduce their working hours further and to stop working completely. OLS estimation would generate bias and a Tobit model should be used. The first solution is to estimate for both workers and non-workers a Type-I Tobit model that accounts for the left-censoring nature of the working hours, which is zero for non-workers. Nonetheless, the Type-I Tobit model cannot explain a situation where the effect of a covariate differs in predicting the probability of working and on the working hours of current workers. Therefore, we adopt a Type-II Tobit model which estimates the working probability and hours of working separately in a bivariate model with the same set of covariates but different estimates of the covariates, and allows the two unobserved individual heterogeneities to correlate with each other. Hours of working for the current workers are estimated using a linear model, and a Probit model is used to model the probability of working.

Similar to a control function approach (Wooldridge, 2015), we include the NRSP participation equation into the system and allow its unobserved effects to correlate with those of working probability and of working hours. The final model is a trivariate error-correlated model where the unobserved heterogeneities of the three endogenous variables are allowed to correlate with each other to account for unobserved effects that determine both individual decisions to participate in the NRSP and their labour supply decisions.

We interact participation status with age group dummies to study the age-specific differences in labour supply response to shocks in pension wealth, and pay special attention to the pension-eligible age of 60. We also separately estimate the model for males and females due to their gender-specific roles in rural families in China. Females spend less time working and more time in informal labour such as doing housework and caring for children or parents than males do. They also suffer from lower labour incomes and rely more on adult children for old-age support (Li, Wang and Zhao, 2018; Shu, 2018), and are subject to worse health status. The descriptive statistics in Table 1.3 also provides observable evidence on the income and health inequality between men and women.

As participating in the pension scheme generates a positive shock on the expected

SSW, any significant negative effect on the labour supply of participants should provide evidence of a crowding-out effect of pension income on labour earnings. We also look at heterogeneous responses by job sectors, education level and asset level. The pension effect is expected to be higher among the lower-wealth, lower educational groups and older agricultural workers.

For comparison between participants and the non-exposed, we use the instrument of community-level placement of the NRSP in 2011. In the first wave of 2011, 256 out of the 454 villages or communities, or 56.39% of the sample did not start the NRSP. These villages or communities all introduced the NRSP by the end of 2012, before the following surveys conducted in 2013 and 2015.

There is a concern that the timing of the NRSP might correlate with some unobserved village- or community-level characteristics that also directly relate to individual labour supply outcomes, and thus confound the estimation of the ITT effect. To mitigate the concern, we use the 2011 community survey data and regress the community-level placement of the NRSP on a rich set of community-level characteristics. They include demographic factors, economic factors, and local labour market conditions. An insignificant correlation would to a certain extent verify the validity of the instrument, which affects individual labour supply only through its effect on predicting individual participation in the NRSP.

Although a qualitative survey of Zhang and Tang (2008) suggests that the choice of pilot communities that started the NRSP as early as in 2008 is related to local population size, education level of residents etc., there is no evidence that the nation-wide expansion of the pension programme is selective. Huang and Zhang (2021) carry out prefecture-level pre-trend tests on a series of local macroeconomic indices and find no evidence of any significant unparalleled trend across the counties. Chen, Wang and Busch (2019) use the survey data of China Family Panel Study (CFPS) and regress the county-level duration of the NRSP programme on a rich set of county-level demographic and economic factors, as well as public health facilities and party secretaries in the county. They also do not find any significant association.

Table 1.3 looks at the rural communities in our sample and reports the pair-wise correlations between the community-level placement of the NRSP and some community-level characteristics from the 2011 community survey. Price level is constructed as an average of the decile rankings of local prices of pork, eggs, rice, and electricity. Other community-level socio-economic factors include local per capita net income, percentage of local population that are above 16; are above 65; work in local enterprises; have a high school degree; are illiterate; have a local *hukou* (residence permit); and work outside the community. It turns out that except for the proportion of illiterate residents, the timing of the NRSP does not seem to relate to any community-level characteristics and can be assumed to be random, as suggested by the literature. The positive correlation between the proportion of illiterate residents and the probability of having the NRSP in place does not support the hypothesis that more economically developed or more educated communities tend to start the NRSP earlier.

Table 1.3: Correlation between instrument and community-level characteristics

Community observables	correlation with NRSP in 2011	p-value
price level	0.057	0.386
income per capita	0.032	0.630
% above 65	-0.099	0.133
% above 16	-0.058	0.378
% employed	-0.033	0.627
% high-school	-0.053	0.428
% illiterate	0.145**	0.029
% locals	0.087	0.187
% out-migrant	-0.023	0.723

Notes: Sample includes the 239 rural communities out of the 454 sample communities in CHARLS 2011 survey.

Cheng et al. (2018*a*) suggest that rural workers might migrate to other counties to join the pension scheme. Given the difficulty of altering registered locations in China, and the short time span that the NRSP took to reach national-wide coverage in 3 years, we assume that this concern is negligible. Nonetheless, Li, Wang and Zhao (2018) drop 358 individuals living in places other than their registered location, or migrant workers, because they are less likely to join the pension programme.

1.5.1 Participation Decision and labour Supply Decision

There are two main labour supply decisions of whether to participate in the labour market and how many hours to work for if entering the labour market. The two decisions are likely determined by different processes. A Type-I Tobit model of working hours cannot explain a situation where the marginal effect of a covariate on the probability of working differs from its effect on reducing hours of working. Therefore, we estimate an instrumental variable Type-II Tobit model with one endogenous variable and specify the working decision and working hours decision using separate distribution assumptions. In practice, we estimate a trivariate error-correlated model and assume the error terms share a trivariate joint normal distribution. The model is estimated by simulated likelihood method of Geweke–Hajivassiliou–Keane algorithm using the conditional mixed process (`cmp`) command in `Stata` (Roodman, 2011), and the standard errors are clustered by individuals. Estimation is carried out separately for males and females to study gender-specific labour supply adjustments after the pension programme.

The following specification is inspired by Semykina and Wooldridge (2018) who study binary response panel data models with sample selection and self-selection problem. Consider a linear model of working hours with endogenous treatment $NRSP_{ijt}$ and a non-random sample selection issue.

$$H_{ijt} = \beta_0 + NRSP_{ijt}\beta_1 + X_{ijt}\beta_2 + \overline{X_{ij}}\beta_3 + v_{ijt3} \tag{1.7}$$

Where H_{ijt} is an observed, continuous variable of weekly working hours. $NRSP_{ijt}$ is the possibly endogenous, pension participation status of individual i in community j in year t. To study age-specific effects of the NRSP participation, we interact $NRSP_{ijt}$ with a set of age group dummies in empirical estimation. X_{ijt} are some observable variables, including individual, household, and community characteristics. $\overline{X_{ij}}$ include the individual-level means over time of those strictly exogeneous and time-varying variables among X_{ijt}, and are included to account for long-run trend of X_{ijt}. v_{ijt3} is the idiosyncratic error term.

We report the average marginal effect of β, which is the effect of a one-unit increase in a given explanatory variable on the expected value of H_{ijt}.

We model the working probability W_{ijt} using a Probit model:

$$W_{ijt} = 1[W^*_{ijt} = a_{i2} + NRSP_{ijt}\delta_1 + X_{ijt}\delta_2 + \overline{X_{ij}}\delta_3 + v_{ijt2} > 0] \tag{1.8}$$

Where W_{ijt} is an indicator of working status that equals 1 if individual i in community j report working for more than 2 hours per week in year t, and equals 0 if he works for only 2 hours or less per week. W^*_{ijt} is a latent variable. The working hour H_{ijt} is only observed for current workers, or when $W_{ijt} = 1$, and is missing if $W_{ijt} = 0$. Again, we interact $NRSP_{ijt}$ with a set of age group dummies in empirical estimation to study a life-cycle pattern of labour supply responses.

We report an average marginal effect of δ, which is the effect of a one-unit increase in a given explanatory variable on the probability of working ($W_{ijt} = 1$) averaged over the sample individuals.

The pension participation decision is also estimated using a Probit model:

$$NRSP_{ijt} = 1[NRSP^*_{ijt} = a_{i1} + NRSP_{jt}\theta_1 + X_{ijt}\theta_2 + \overline{X_{ij}}\theta_3 + v_{ijt1} > 0] \tag{1.9}$$

Where the individual pension participation status is a function of $NRSP_{jt}$, an indicator of the placement of the NRSP in community j in year t. The $NRSP_{jt}$ serves as an instrument that satisfies both a monotonicity and an exclusion restriction, and thus enables the estimation of the ATE.

Assume that the regression errors $(v_{ij1}, v_{ij2}, v_{ij3})$ share a zero-mean trivariate joint normal distribution. We report the pairwise correlations between unobserved effects, $\rho_{NRSP-hrs}$, $\rho_{work-hrs}$ and $\rho_{NRSP-work}$ at the bottom of the estimation tables. The correlation can only be estimated from the non-missing data. A non-zero correlation between v_{ij3} and v_{ij2}, or the error term of the working hours equation and that of the working probability equation indicates a significant correlation between the unobservables that predict hours of working and those predicting the probability of working, conditioning on the observables. If $\rho_{work-hrs}$ is positive, it means that workers who work for more hours have higher incentives to stay in the labour force than workers who work for fewer hours. These unobserved effects can be motivation, working incentives such as financial difficulties. Estimating the truncated working hours or working probability separately would suffer from the sample selection bias. A significant non-zero $\rho_{NRSP-work}$ or $\rho_{NRSP-hrs}$ suggests that conditioning on the observables, individual decision to participate in the pension programme is correlated with individual labour participation decision or labour supply decision via some unobserved effects. People who are more likely to participate in the NRSP may be those planning to stop working and or working partially. Omitting the sample selection issue and self-selection into the pension programme would cause bias in estimating the pension effect. The size and direction of the bias depend on the magnitudes and signs of the correlations between unobserved effects that determine pension participation and those determining labour supply.

1.5.2 Contribution Decision

Apart from the decision to participate in the NRSP, the rural age-ineligible also need to consider how much to contribute to individual pension accounts every year after their enrolment. In this section, we focus on the subsample of the age-ineligible rural residents and study how their choices of annual contribution levels affect their labour supply behaviours. Because more than 70% of NRSP participants in our sample contribute at the minimum requirement level of 100 RMB per year, predicting the levels of annual contribution using a linear model can violate the normal distribution assumption of the linear model. Moreover, individuals contributing at a level higher than the minimum requirement may differ systemically from those meeting the minimum requirement in terms of commitment to the pension scheme and willingness to work. Therefore, instead of using the amount of annual contribution, we construct a binary indicator of high contributors that assigns 1 for NRSP participants who contribute at a level higher than the minimum requirement, and equals 0 for participants who contribute at the minimum requirement. Shu (2018) studies the heterogeneous effects of NRSP participation on labour supply behaviour for the two groups of high and low contributors, while our method aims to study whether moving from low to high contributions is associated with a change in labour supply behaviour.

Because contribution levels are only observed among participants in the NRSP, there is a potential selection bias that individuals more willing to contribute to their pensions are more likely to participate. Similar to the way we estimate the hours of working and the probability of working, we estimate the pension participation and contribution decisions together using a bivariate Probit model. To identify the probability of being a high contributor, we use community-level, heterogeneous years of introducing NRSP and construct the community-level duration of the NRSP programme as the difference between the survey year and the year when the NRSP started in local communities. Previous studies have used county-level duration of the NRSP programme as an instrument for individual participation probability (Chen, Wang and Busch, 2019; Cheng et al., 2018*a*,*b*; Shu, 2018). The assumption is that the longer the programme has been in place, the more likely people will participate and will contribute at a higher level to their pensions after a build-up of knowledge about the programme and trust in the local governments. Again, we assume that the placement of the NRSP is random and will not directly affect individual behaviours.

To study the effect of moving from low to high contributors on labour supply behaviour, we estimate the following multivariate error-correlated model that is built up on the trivariate model in the previous section. Specifically, we include an extra regression that predicts the probability of contributing at a level above the minimum requirement for the NRSP participants:

$$HC_{ijt} = 1[HC^*_{ijt} = a_{i4} + NRSP\_DUR_{jt}\varphi_1 + X_{ijt}\varphi_2 + \overline{X_{ij}}\varphi_3 + v_{ijt4} > 0] \tag{1.10}$$

Where HC_{ijt} is an indicator of high contributors that equals 1 if individual i in community j contributes to the NRSP at a level higher than the local minimum requirement in year t, and equals 0 if contributes at the local minimum requirement. HC_{ijt} is only observed for age-ineligible participants in the NRSP, and is missing for non-participants. $NRSP\_DUR_{jt}$ includes three indicators of the duration of the NRSP in community j in survey year t. $NRSP \leq 0y$ equals 1 for communities that had not started the NRSP in 2011 or had just started it earlier that year, and equals 0 for the rest in 2011 or for all communities in 2013 or 2015, when the NRSP had reached nation-wide coverage. $NRSP(0, 2y]$ equals 1 for communities that had started the NRSP and had it in place for 2 years or less in the survey year. $NRSP(2, 4y]$ equals 1 for communities that had had the NRSP in place for more than 2 years but less than or equal to 4 years in the survey year. $NRSP > 4y$ equals 1 for communities that had had the NRSP in place for more than 4 years in the survey year.

$NRSP\_DUR_{jt}$ serves as an instrument that satisfies both a monotonicity and an exclusion restriction, and thus enables the estimation of the ATE. Similarly, we assume the error terms share a multivariate normal distribution. The hypothesis is that the higher the contribution, the higher the expected social security wealth, and the stronger is the crowding-out effect of pension income on labour earnings and labour supply.

1.6 Data and Summary Statistics

The paper uses a nation-wide, biennial household survey data from the China Health and Retirement Longitudinal Studies (CHARLS). The baseline survey started in 2011 and covered 17,708 individuals who are aged 45 and above and come from 10,257 households across 450 villages or communities in 150 counties, and 28 provinces in China[4]. These respondents are followed up in 2013, 2015 and 2018, though some new participants are included to replace the drop-outs. CHARLS is designed to be comparable with the Health and Retirement Survey (HRS) in the US and similar surveys around the world. The paper uses data from the three waves of surveys collected in 2011, 2013, and 2015.

We choose the CHARLS because it contains the most detailed and comprehensive information on older people's retirement and pensions, as well as health status. The middle-aged and the elder that the survey focuses on are ideal for studying the relationship between exit from the labour force and pension. Another two nation-wide, longitudinal surveys in China, the China Family Panel Studies (CFPS) and the Chinese Longitudinal Healthy Longevity Survey (CLHLS) are also biannual surveys and do not provide as much information about labour supply behaviours and pension as the CHARLS does[5][6].

The whole sample is restricted to individuals that: 1) are registered rural residents (holding rural-hukous; 2) have not formally retired; 3) do not miss information on pension participation and labour supply. In other words, we focus on the target group of the NRSP that are eligible to participate in the scheme once it is introduced in the local community.

1.6.1 Demographic Characteristics

To account for factors that affect both pension participation and labour supply decisions of individuals, we control for a large set of SES variables, time, and location dummies. Individual-level variables contain age (rage) and age squared (rage_sq) divided by 100, indicators for age groups (raged5055, raged5560, rabove60), indicators for educational level (rnprimary, rprimary, rsecondary, rhighabove), indicators for marital status (rmarried). Average ages within the rural, age-eligible group and the age-ineligible group are stable across waves, due to the inclusion of new and younger

[4] Tibet is not included in the survey, and two other provinces of Hainan and Ningxia are also not represented in the study due to their relatively small sizes of population

[5] The CFPS collects information of all household members of the selected households and does not have information about individual NRSP participation status for some waves.

[6] The CLHLS collects information about elder people aged over 65, with an average age of 84 and would not be suitable for studying older workers' labour market transitions as most of the respondents are out of the labour force.

respondents in follow-up waves. Marriage ratio is overall very high, above 95% among the age-ineligible group, and above 85% among the age-eligible males, above 70% among age-eligible females. The lower marriage ratio among the elder group is mainly due to widowhood. In terms of educational level, the younger group generally holds higher degrees than the older group does, and men are more educated than women.

Household-level variables include the number of children (hchild), the number of household members (hhhres), and the log of the value of household durable assets (lnhhadurbl)[7]. We use the value of household durable assets to measure the household resource because it is a better measure of long-run resources and has fewer missing values [8]. Measures of household income and expenditure in the survey show greater variation and have more missing data.

We also control for where the respondents are located using indicators of living in urban areas (urban_nbs) and of seven area types (area1-7) defined by National Statistics Bureau. Among them, urban areas include city-centres, city-towns, small city-centres, small city-towns and special areas, and rural areas include towns and villages. Provincial differences can play an important role in the level of health inequality and health–SES gradient because of health care and other prices, inherent healthiness of the area, public health infrastructure, and other factors (Lei et al., 2014*b*). Therefore, we control for provincial differences in health and labour market conditions by including 28 provinces' dummies.

Finally, indicators of the survey years of 2013 (y2013) and 2015 (y2015) are also included to account for the time trends of aggregate health and labour supply behaviour, time-varying reporting changes and effects of age that are not captured by age and its quadratic term.

Table 1.4 summarizes the demographic characteristics by the pension-eligible age of 60, gender, and survey year. Participation rate jumps from 2011 to 2013 and stabilizes from 2013 to 2015, because about 50% of the sampled communities do not have the NRSP in 2011 and by the end of the 2012, all of them have implemented the scheme. Participation rate is higher among the age-eligible than the age-ineligible, and are similar for men and women. Working probability, probability of doing non-agricultural work and hours of work are higher among the younger cohorts and men.

The mean age is 68 for age-eligible men and women, around 52 for age-ineligible men and women. Educational achievement is higher for men in the younger cohorts,

[7]Durable assets include cameras, air conditioners, mobile phones, furniture, music instruments, valuable decorations and ornaments, treasures and precious metals, antiques or paintings or artistics work, refrigerator, washing machine, TV, computer, stereo system, video camera. Respondents are first asked whether they own each type of these assets, and then the current values of these assets. The value of household durable assets sums up the values of all assets respondents own.

[8]It might be the case that as long as respondents report the values of any one of the durable assets listed above (they might well miss reporting some of the items they own), the variable will not be missing.

Table 1.4: Descriptive statistics of individual and household characteristics

	rural age-eligible						rural age-ineligible					
	Male			Female			Male			Female		
	2011	2013	2015	2011	2013	2015	2011	2013	2015	2011	2013	2015
NRSP_{ijt}	0.28	0.72	0.73	0.29	0.72	0.71	0.27	0.66	0.63	0.28	0.67	0.64
Pension_{ijt}	0.69	0.88	0.88	0.69	0.88	0.89						
contribution							193.5	189.9	224.2	212.3	173.5	214.9
							(295.4)	(423.6)	(528.8)	(420.2)	(464.0)	(593.3)
work	0.71	0.70	0.70	0.54	0.56	0.55	0.93	0.92	0.93	0.84	0.82	0.82
farmwork	0.60	0.57	0.53	0.50	0.51	0.47	0.55	0.51	0.41	0.66	0.63	0.55
nfarmwork	0.11	0.13	0.17	0.05	0.06	0.08	0.38	0.42	0.52	0.19	0.20	0.28
hours	40.3	39.6	40.2	34.1	33.5	32.5	49.1	49.2	49.1	42.9	42.3	43.6
	(25.6)	(24.7)	(26.3)	(24.9)	(24.7)	(24.4)	(24.1)	(24.2)	(23.4)	(26.0)	(26.3)	(26.7)
farmhours	31.7	30.6	27.2	29.9	29.5	25.9	26.4	24.7	19.1	30.8	29.4	24.7
	(26.5)	(25.7)	(25.6)	(25.1)	(24.7)	(23.5)	(29.0)	(28.9)	(27.0)	(27.7)	(27.6)	(26.9)
nfarmhours	8.6	9.0	13.0	4.2	4.1	6.7	22.7	24.5	30.0	12.1	12.9	18.9
	(22.4)	(22.4)	(27.2)	(15.5)	(15.5)	(19.7)	(30.6)	(31.0)	(31.1)	(25.5)	(26.1)	(30.5)
age	67.9	68.1	68.3	68.4	68.5	68.7	52.2	52.3	51.4	51.3	51.4	50.6
	(67.9)	(68.1)	(68.3)	(68.4)	(68.5)	(68.7)	(52.2)	(52.3)	(51.4)	(51.3)	(51.4)	(50.6)
primary	0.33	0.32	0.32	0.13	0.13	0.14	0.24	0.25	0.40	0.19	0.21	0.33
secondary	0.12	0.13	0.15	0.03	0.04	0.05	0.36	0.37	0.30	0.18	0.21	0.18
highabove	0.02	0.03	0.04	0.00	0.00	0.01	0.14	0.15	0.12	0.05	0.05	0.04
married	0.84	0.84	0.84	0.69	0.71	0.71	0.95	0.95	0.95	0.95	0.95	0.95
numchild	3.3	3.3	3.2	3.6	3.6	3.6	2.2	2.2	2.2	2.3	2.3	2.3
	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)	(1.6)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)	(1.0)
hhres	3.5	3.6	2.9	3.4	3.5	2.8	3.9	3.9	3.4	4.0	4.0	3.4
	(2.0)	(2.0)	(1.4)	(2.0)	(2.0)	(1.4)	(1.7)	(1.7)	(1.3)	(1.7)	(1.7)	(1.3)
lnhhadurbl	6.1	6.4	6.4	6.0	6.2	6.2	7.2	7.4	7.6	7.2	7.4	7.5
	(2.4)	(2.4)	(2.3)	(2.5)	(2.5)	(2.5)	(1.6)	(1.8)	(1.7)	(1.5)	(1.8)	(1.7)
urban_nbs	0.21	0.22	0.23	0.25	0.25	0.26	0.26	0.26	0.29	0.27	0.26	0.28
N	2,588	2,903	3,148	2,815	3,295	3,597	3,538	3,312	3,620	4,123	3,978	4,126

followed by age-ineligible women, age-eligible men and lowest for age-eligible women. Marriage ratio is lower for the older cohorts due to widowhood. The younger cohorts have less children than the older cohorts do due to family planning policy, and report higher level of household durable assets' value.

1.6.2 Pension Participation and Labour Supply Status

The indicator of participants is constructed based on self-reported participation status, and pensioners are identified in a similar way. We assign 0 to non-participants and contributors and 1 to pensioners for the pensioner indicator.

As for the outcome variables, the participation ratio more than doubles from 2011 to 2013, due to the rapid expansion of the NRSP across the nation, and stabilizes after 2013. Above 70% of the age-eligible and above 60% of the age-ineligible join the pension programme after it was introduced in their residential communities. Not all age-eligible participants receive pensions as the ratio of pensioners is lower than the ratio of participants. Although following an uprising trend, the annual contribution

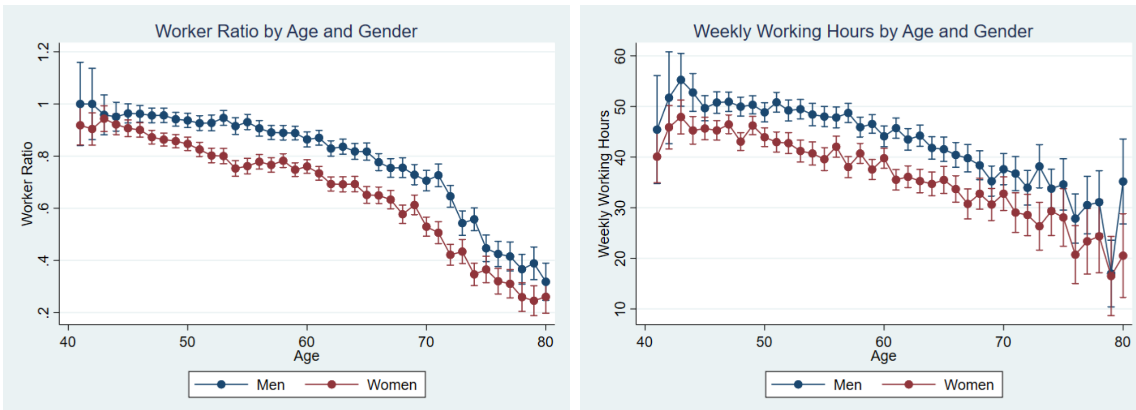

Figure 1.1: Labour Supply by Gender and Age Groups

Figure 1.2: NRSP Participation and Receipt by Gender and Age Groups

remains at a low level of around 200 RMB per year. The minimum requirement is 100 RMB in all places except for Beijing, Tianjin, and Shanghai where the level of minimum requirement is higher.

The self-reported working indicator assigns 1 to respondents who report to have engaged in any work, either farming- or non-farming work, in the previous year. It assigns 0 to respondents who report to be unemployed, retired, or have never ever worked before. There are non-substantial amounts of respondents who reported being retired in earlier waves going back to the labour force in later waves, especially for those retiring from agricultural jobs. We regard it more accurate to define these people as being economically inactive or not working, instead of being retired. The major difference is that retirees receive retirement pensions and not usually go back to work after retirement, while the economically inactive may receive no or little pensions and are more likely to go back to work for example when they are faced with financial stress.

Indicators of farming workers or non-farming workers are constructed based on the type of main jobs they take[9]. They are assigned 0 for the economically inactive and workers in other sectors.

We also construct a continuous variable of weekly working hours in main jobs based on the reported number of hours worked each day and the number of days worked each week in the previous year. We focus on workers' labour supply behaviour in their main jobs because less than 1% of the respondents report to have a side job. Figure 1.1 provides graphical illustrations of the proportion of workers and the distribution of weekly working hours by gender and age group. There is a trend of declining working probability when people grow old, and the decline is steeper for women than for men. The hours of working or labour intensity also declines with age, similarly for men and women. Women are less likely to work and work less hours compared to men across all age groups. Figure 1.2 shows the ratio of NRSP participants and pensioners across age groups. Participation rate is among the age-eligible groups, who are almost certain to receive pensions when they enrol in the programme.

Several facts about the labour supply outcomes are observable based on Table 1.4. Firstly, the employment rates are about 70% and 55% among age-eligible men and women, respectively, but are as high as 93% and 82% among age-ineligible men and women. The majority of workers take farming jobs. Only about 13% of age-eligible men and 5 to 8% of age-eligible women take non-farming jobs. The ratio is higher among the age-ineligible, as above 30% of men and above 20% of women do non-farming work. There is a trend of transition from farm-working to non-farm

[9] The respondent is coded as doing farm-work if he reports working (1) in agricultural work (including farming, forestry, fishing, and husbandry for your own family or others) for more than 10 days in the past year, (2) for other farmers for wages for at least ten days in the past year, or (3) for their own household for at least ten days in the past year.

working, more observable among the younger cohorts. The working probability and weekly working hours do not change much over time for both the age-eligible and the age-ineligible.

1.7 Empirical Results

By using community-level placement of the NRSP in a given year as the instrument, we compare participants who enrolled in the NRSP once it was introduced in the local community, with people coming from communities that did not have the NRSP, given that they share similar observable and unobservable characteristics.

Table 1.5 compares the key outcomes for the group of communities that did not have the NRSP in 2011 but had it in 2013, and the group of communities that had the NRSP in 2011.

Table 1.5: Pension status and labour supply by treated and comparison groups

	rural age-ineligible						rural age-eligible					
	Male			Female			Male			Female		
	2011	2013	2015	2011	2013	2015	2011	2013	2015	2011	2013	2015
Had NRSP by 2011												
NRSP partcipants(=1)	0.48	0.62	0.60	0.51	0.65	0.60	0.46	0.62	0.64	0.51	0.69	0.69
work(=1)	0.86	0.86	0.90	0.73	0.71	0.76	0.63	0.61	0.62	0.45	0.48	0.48
annual working hours	1900	1897	1777	1673	1613	1595	1602	1478	1437	1372	1254	1172
farm work(=1)	0.46	0.42	0.38	0.54	0.51	0.49	0.50	0.48	0.46	0.40	0.43	0.42
farm work hours	1558	1478	1297	1432	1309	1146	1426	1285	1162	1300	1138	1014
nonfarm work(=1)	0.28	0.30	0.38	0.12	0.13	0.19	0.09	0.08	0.12	0.02	0.03	0.05
nonfarm work hours	2234	2253	2093	2213	2235	2332	2213	2334	2306	1457	2338	1764
Had NRSP by 2013												
NRSP partcipants(=1)	0.08	0.51	0.49	0.09	0.51	0.48	0.06	0.47	0.47	0.09	0.53	0.51
work(=1)	0.82	0.85	0.88	0.65	0.69	0.70	0.50	0.51	0.51	0.36	0.40	0.38
annual working hours	1937	1938	1932	1791	1788	1778	1679	1590	1541	1468	1422	1319
farm work(=1)	0.39	0.37	0.29	0.44	0.44	0.37	0.40	0.39	0.36	0.32	0.35	0.32
farm work hours	1553	1483	1339	1489	1416	1247	1507	1389	1240	1364	1320	1145
nonfarm work(=1)	0.30	0.34	0.42	0.14	0.16	0.22	0.07	0.07	0.10	0.02	0.02	0.04
nonfarm work hours	2222	2174	2140	2177	2199	2152	2358	2103	2271	2018	1605	1886

A large discrepancy exists between the two groups in terms of the ratio of NRSP participants and pensioners in 2011, making the community-level placement of the NRSP a potentially valid instrument to predict individual, actual participation in the NRSP. Although working probabilities appear to be similar across the two groups, farm working probability is slightly lower in communities that started the NRSP after 2011. Communities that had the NRSP in 2011 overall show a lower level of annual working hours and annual farm working hours. A potential explanation is that people started to reduce farm-working or total working hours or even stop farm working completely after participating in the NRSP. Nonetheless, we need to estimate the effect of pension participation empirically using causal inference methods to verify the hypothesis.

Before estimating the trivariate model specified in the previous section, we estimate the simplified multivariate model in which no endogenous variable of NRSP participation status or contribution levels enters into the model of working probability or working hour. Error terms are still assumed to share a multivariate normal distribution and we estimate their pairwise correlations to provide evidence on the existence of unobserved effects that predict both pension participation/ contribution, and labour supply, after controlling for some observable characteristics. Specifically for Table 1.6, we control for individual- and household-level variables and the individual-level means for the time-varying variables among them, survey year dummies and provincial dummies.

As shown in Table 1.6, columns 1 and 3 estimate jointly the probability of participating in the NRSP, the probability of working, and the hours of work if currently in the labour force using the sample of non-formally retired, rural residents. Columns 2 and 4 estimate a system of four equations, adding on the above three-equations model an extra function estimating the probability of contributing to the NRSP at a level higher than the minimum requirement for the NRSP participants. The model is to study whether being a high contributor, or contributing at a level higher than the minimum requirement, is related to any further change in individual labour supply behaviour. We restrict the sample to non-formally retired rural residents who are aged below 60 and need to pay for their rural pension accounts if they participate in the NRSP. Probit models are used for binary dependent variables including working or not, participating in the NRSP or not, contributing above or at the minimum requirement level. A linear model is used to estimate hours of working. Standard errors are clustered by individual levels.

The pairwise correlations of unobserved effects reported in columns 1 and 3 suggest that for both males and females, the individual decision to participate in the NRSP is endogenous to retirement decision ($\rho_{NRSP-work}$ is significant in all columns), and unobserved factors that predict participation in the NRSP also predict staying in the labour force till older age. They can be financial difficulties, disadvantaged socio-economic conditions that drive rural workers to stay in the labour force until very old age. Unobserved factors that predict staying in the labour force also predict less hours of working for women, as shown by the significant and negative $\rho_{work-hrs}$ in column 3. Women staying in the labour force until old age are likely to gradually reduce their hours of working, while the pattern is not significant for men, as men tend to work until older ages compared to women.

The pairwise correlations of unobserved effects reported in column 4 show that among age-ineligible women, the unobserved effects that predict the probability of

Table 1.6: Endogeneity of NRSP participation decision and contribution decision

	Males		Females	
	(1)	(2)	(3)	(4)
Correlation of errors:				
$ln\sigma_{hrs}$	3.177***	3.157***	3.230***	3.245***
	(0.006)	(0.018)	(0.007)	(0.007)
$\rho_{NRSP-work}$	0.080***	0.093***	0.065***	0.078***
	(0.017)	(0.027)	(0.014)	(0.019)
$\rho_{NRSP-hrs}$	-0.003	-0.001	-0.000	0.000
	(0.012)	(0.016)	(0.012)	(0.015)
$\rho_{work-hrs}$	-0.001	-0.210	-0.151**	-0.039
	(0.072)	(0.391)	(0.074)	(0.103)
$\rho_{NRSP-contribution}$		1.568***		-0.509***
		(0.280)		(0.157)
$\rho_{work-contribution}$		0.090**		-0.020
		(0.039)		(0.031)
$\rho_{hrs-contribution}$		0.033		0.050**
		(0.020)		(0.022)
Observations	18828	10386	21508	12189
log likelihood	-86303	-52932	-90902	-59830

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in the individual level. The sample is restricted to rural residents not formally retired and not missing in the outcome variables in columns 1 and 3, and further to those aged below 60 in columns 2 and 4.

contributing above the minimum requirement level are significantly and negatively related to the unobserved effects that predict the probability of participating in the NRSP. This suggests that the majority of female participants contribute at the minimum level instead of above the requirement. In contrast, the positive and significant $\rho_{NRSP-contribution}$ in column 2 suggests that age-ineligible men who join the NRSP tend to contribute at a level higher than the minimum requirement. In other words, among the age-ineligible men, only those who have high commitment to the scheme and willing to contribute more to their pensions join the NRSP. For women, high contributors are significantly more likely to work for longer hours, as shown by the significant and positive $\rho_{hrs-contribution}$ in column 4. The significant and positive $\rho_{work-contribution}$ in column 2 suggests that for men, high contributors are more likely to work than low contributors do. Because there is no unobserved confounder that predict both hours of work and working probability for the age-ineligible men or women (the $\rho_{work-hrs}$ is not significant in columns 2 and 4), in studying labour supply behaviours of age-ineligible rural residents, we estimate a two separate trivariate models of the probability of participating in the NRSP, the probability of contributing above the minimum level and the probability of working or hours of work for the age-ineligible groups. By doing so, we improve the efficiency of the model and estimate the correct standard errors compared to the two-step IV estimation.

1.7.1 NRSP Participation Decision

Before we study the effect of NRSP participation on individual labour supply behaviours, we explore factors that predict the probability of enrolling in the pension scheme. Table 1.7 compares the estimation results using different specifications for the NRSP participation equation (1.9), separately for men and women. Specifically, the baseline models (columns 1 and 5) only control for individual and household characteristics, survey year dummies and dummies of boarder administrative areas. Columns 2 and 6 add on to columns 1 and 5 extra covariates of the provincial dummies. Columns 3 and 7 add to columns 2 and 6 individual-level, over-time means of time-varying exogenous variables, including marital status, number of living children, number of household members, and the logged values of household durable assets. Although the introduction time of the NRSP is random, communities vary in local implementation and the amount of basic pension benefits due to their heterogeneous socio-economic conditions. Columns 4 and 8 add to columns 3 and 7 some community characteristics, including the local price level constructed as an average of the decile rankings of local prices of pork, eggs, rice, and electricity, per capita local income levels, the proportion of local population that are aged above 65, the proportion of local population that hold high-school degrees, the proportion of local population

Table 1.7: Determinants of Individual Participation in the NRSP

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NRSP in place	1.742***	1.911***	1.913***	1.913***	1.760***	1.927***	1.929***	1.948***
	(0.047)	(0.053)	(0.053)	(0.057)	(0.044)	(0.049)	(0.048)	(0.053)
rage	0.109***	0.110***	0.108***	0.102***	0.094***	0.094***	0.093***	0.093***
	(0.012)	(0.012)	(0.012)	(0.014)	(0.010)	(0.009)	(0.010)	(0.010)
rage_sq	-0.085***	-0.082***	-0.080***	-0.075***	-0.072***	-0.070***	-0.068***	-0.068***
	(0.010)	(0.010)	(0.010)	(0.011)	(0.008)	(0.008)	(0.008)	(0.008)
rprimary	-0.004	0.030	0.033	0.035	-0.022	0.032	0.032	0.014
	(0.029)	(0.029)	(0.029)	(0.031)	(0.029)	(0.029)	(0.029)	(0.031)
rsecondary	0.069**	0.056*	0.057*	0.043	0.036	0.066*	0.064*	0.074*
	(0.032)	(0.033)	(0.033)	(0.035)	(0.037)	(0.037)	(0.038)	(0.040)
rhighabove	0.007	0.016	0.017	-0.008	0.017	0.061	0.057	0.060
	(0.047)	(0.047)	(0.047)	(0.051)	(0.065)	(0.063)	(0.063)	(0.067)
rmarried	0.016	0.038	0.079	0.104	0.030	0.004	-0.125	-0.147
	(0.040)	(0.040)	(0.117)	(0.123)	(0.033)	(0.033)	(0.096)	(0.106)
hchild	0.048***	0.032***	0.070***	0.077***	0.022**	0.013	0.084***	0.077***
	(0.010)	(0.010)	(0.024)	(0.027)	(0.009)	(0.009)	(0.023)	(0.025)
hhhres	-0.004	0.004	-0.018	-0.020*	-0.008	0.000	-0.005	-0.007
	(0.007)	(0.007)	(0.011)	(0.012)	(0.006)	(0.006)	(0.010)	(0.011)
lnhhadurbl	0.012**	0.015***	0.016**	0.014	0.015***	0.014***	0.011	0.010
	(0.006)	(0.006)	(0.008)	(0.009)	(0.005)	(0.005)	(0.007)	(0.008)
y2013	0.488***	0.451***	0.448***	0.483***	0.466***	0.437***	0.428***	0.462***
	(0.027)	(0.029)	(0.029)	(0.031)	(0.025)	(0.027)	(0.027)	(0.029)
y2015	0.464***	0.432***	0.414***	0.437***	0.419***	0.390***	0.365***	0.378***
	(0.027)	(0.029)	(0.029)	(0.032)	(0.025)	(0.027)	(0.028)	(0.030)
urban_nbs	-0.227***	-0.174**	-0.173**	0.011	-0.140**	-0.139**	-0.141**	0.047
	(0.075)	(0.081)	(0.081)	(0.106)	(0.064)	(0.067)	(0.067)	(0.093)
_cons	-5.302***	-5.505***	-5.466***	-5.098***	-4.788***	-4.843***	-4.823***	-4.773***
	(0.374)	(0.398)	(0.399)	(0.452)	(0.295)	(0.315)	(0.317)	(0.353)
Provincial Dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$\overline{X_{ij}}$	No	No	Yes	Yes	No	No	Yes	Yes
Community Controls	No	No	No	Yes	No	No	No	Yes
$\rho_{NRSP-work}$	0.077	0.036	0.030	0.064	0.113***	0.101***	0.104***	0.107***
	(0.059)	(0.051)	(0.051)	(0.054)	(0.041)	(0.038)	(0.038)	(0.041)
$\rho_{NRSP-hrs}$	-0.088**	-0.068**	-0.068**	-0.094***	-0.054*	0.006	0.006	-0.023
	(0.035)	(0.031)	(0.032)	(0.034)	(0.030)	(0.028)	(0.028)	(0.030)
Observations	18636	18635	18635	16018	21227	21227	21227	18330
log likelihood	-86489	-85601	-85569	-73576	-91174	-90021	-90001	-78014

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for time-varying variables, community-level characteristics. The sample is restricted to rural residents not formally retired and not missing in the outcome variables.

that hold rural residential permits (hukous), and the dependency ratio calculated as the number of local residents aged above 65 dividing by the number of local residents aged above 16. Among them, the proportion of residents that hold rural residential permits can distinguish urban from rural communities. The community characteristics are collected from the 2011 community survey, the only publicly available community survey from CHARLS. We treat these community characteristics as time-invariant variables, which might not be a serious problem given the fact that these characteristics are relative stable within four years of time span.

Placement of the pension scheme in the local community is the most significant factor that predicts individual participation in the NRSP, followed by time effect and age effect, suggesting that it is a valid instrument to identify the effect of pension participation on individual labour supply. The significant effects of survey year indicators suggest that over time, eligible individuals are more likely to enrol in the scheme after a build-up of knowledge about it and trust in local governments. It can also be the case that individuals wait and join the scheme when they are older and approaching the pension-eligible age.

Older people are more likely to join the pension scheme than the younger cohorts, mainly due to the fact that those aged above 60 when the pension started are eligible to receive basic benefits without making any contributions. Conditions on annual contribution can discourage the younger cohorts from joining the scheme. Number of children can be another proxy for cohort effects, as one-child policy introduced in the late 1970s affects younger cohorts more than the older cohorts, and the younger cohorts tend to have less children compared to the older cohorts. Educational levels are positively related to the probability of participating in the scheme as more educated people have a lower time preference and are more likely to save for their pensions than the low-educated, poorer people do, though the effect is marginally significant. Note that the illiterate individuals are used as the reference group. Having finished high school or higher education does not increase the chance of enrolment. Rural-residents rarely finished high school in the old days when our sampled individuals are in their schooling age, or they would have migrated to the urban areas and changed into urban-hukous, worked in non-agricultural sectors and join the employees pension program.

Rural workers staying in rural communities are likely to differ systemically from rural workers staying in urban communities and the indicator of living in urban areas ($urban_n bs$) can control for that. Specifically, rural-hukou holders living in urban areas are more likely to be migrant workers taking non-agricultural jobs and might be enrolled in employees pension program. In the case that they are not enrolled in the employees' pension program, they need to go back to their registered residential towns or villages to participate in the NRSP, which reduces their willingness to participate after the scheme was introduced. Other factors such as income levels and living arrangements (migrant workers are less likely to live with family compared to local workers) can also affect migrant workers' probability of participating in the NRSP. The negative effect of living in urban areas confirms that migrant workers are less likely to participate in the NRSP. The effect was picked up by community characteristics that distinguish rural and urban areas in columns 4 and 8.

People reporting higher levels of durable asset value are more likely to enrol in the pension programme. The wealth effect can have a similar pattern as the educational effect in predicting individual enrolment, as very poor individuals are less likely to join the scheme if they have not reached the pension-eligible age when the NRSP starts. It may also be explained by the fact that older people keep more durable assets than younger people do. Significant levels and magnitudes of the estimated coefficients do not change much for most of the reported covariates. Columns 4 and 8 lose some observations due to their missing values of the community-level characteristics. Therefore, columns 3 and 7 that control for both provincial dummies and individual-level means of time-varying variables, but not community-level characteristics are our preferred models, and will be used for the heterogeneity analyses in the following section.

1.7.2 Labour Supply Decision and NRSP Participation Decision

Table 1.8 and 1.9 report the trivariate estimation outputs of the working probability equation (equation 1.8) and the working hours equation (equation 1.7), respectively. We separate the analyses for men and women and compare them with the estimation results using different sets of control variables. We focus on changes in sizes and significance levels of the variables of interest, namely, the interactions of individual NRSP participation status with an exclusive age group dummies among different specifications.

There is an overall lack of significant effect of NRSP participation on men's working probability, except that male participants aged above 75 are more likely to stop working compared to the non-participants. The effect is only marginally significant and appears only in the specification that accounts for community characteristics (column 4). Although non-significant, the signs of participation effects for men are positive across most of the age groups, suggesting that male participants do not stop working completely after joining the pension scheme or receiving basic benefits from the scheme. In contrast, female participants are less likely to work across all age groups, although the effects are only marginally significant for those aged above 50 and those aged above 70, who are either approaching pension-eligible age or are too old to work.

The negative and significant correlations between unobserved effects that predict working probability and those predicting hours of work for the current female workers ($\rho_{work-hrs}$ in columns 5-8) are more observable among agricultural workers, as shown in Tables 1.11 and 1.12. Among female agricultural workers, there are unobserved effects such as job characteristics, caring responsibility, individual savings that affect both their probability of work and hours of work. In contrast, women taking non-agricultural jobs tend to earn more and stop working earlier than those doing agricultural work, though also work for longer hours when they are in the job. The pattern applies for both men and women. Not accounting for the unobserved effects and estimating a single model for the truncated variable of hours of work for current workers can produce biased estimation due to omitting bias.

The significant and positive $\rho_{NRSP-work}$ across all models for women (columns 5-8 in Table 1.8) suggests that the unobservable factors predicting the probability of participating in the NRSP also predict the probability of working. Heterogeneity analyses using population subgroups show that the effect mainly comes from older agricultural workers (column 7 in Table 1.11). These unobserved effects can be financial difficulties, disadvantaged socio-economic conditions that drive female agricultural workers to stay in the labour force until very old age, and to register for the basic benefits from the NRSP.

For men, there is no significant difference in the probability of NRSP participation between current workers and non-workers, although agricultural workers are more likely to participate than non-agricultural workers do, after conditioning on observable characteristics (column 3 in Table 1.11). Not controlling for the endogeneity of NRSP participation that female participants are more likely to be working than female non-participants will downwardly bias the negative labour supply effect of participation in the NRSP. The difference in participation probability between male and female workers is consistent with the hypothesis that the NRSP has the biggest impact on the low-income groups and the groups of workers who have lower productivity such as older female agricultural workers.

Among socio-economic factors, very old age, represented by the rescaled square of age, is related to a lower probability of working and less hours of work. Marriage is related to a higher probability of staying in the labour force. More than 70% of the age-eligible sample and 95% of the age-ineligible sample are married. The unmarried group mainly consists of those entering widowhood. Therefore, marital status also picks up age effects in the sense that it is negatively related to age. Having more children and household members are correlated with earlier exit from the labour force for women and especially men, suggesting that supports provided by adult children are an important determinant of labour market participation among older rural residents.

The time trend of gradually reducing hours of work is only significant for men. Rural residents living in urban areas, represented by *urban_nbs*, stop working earlier than those living in the rural areas, although they also tend to work longer hours when they are working, as shown in Table 1.9. Rural migrant workers living in urban areas are more likely take physically demanding, non-agricultural jobs and work for long hours, and tend to stop working earlier than individuals working in their own farms. People doing family agricultural work or informal work have more discretion on the hours of working at the cost of low job security and compensation. Community-level socioeconomic condition also affects labour supply (not reported in the Tables).

Table 1.8: Effect of NRSP Participation on Working Probability

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
rNRSP_below45	-0.067	-0.017	-0.003	-0.111	0.092	0.119	0.118	0.124
	(0.222)	(0.222)	(0.223)	(0.241)	(0.119)	(0.117)	(0.117)	(0.129)
rNRSP_4549	0.026	0.097	0.113	-0.020	-0.031	-0.007	-0.009	-0.042
	(0.116)	(0.107)	(0.107)	(0.113)	(0.078)	(0.074)	(0.074)	(0.080)
rNRSP_5054	-0.009	0.097	0.114	0.050	-0.154**	-0.117*	-0.122*	-0.110
	(0.107)	(0.097)	(0.098)	(0.104)	(0.075)	(0.071)	(0.071)	(0.076)
rNRSP_5559	-0.004	0.107	0.115	0.037	-0.145**	-0.097	-0.102	-0.138*
	(0.104)	(0.094)	(0.094)	(0.099)	(0.072)	(0.068)	(0.068)	(0.073)
rNRSP_6064	-0.028	0.083	0.091	0.044	-0.054	-0.024	-0.028	-0.028
	(0.101)	(0.091)	(0.091)	(0.096)	(0.071)	(0.067)	(0.067)	(0.072)
rNRSP_6569	-0.041	0.053	0.061	0.004	-0.025	-0.025	-0.029	-0.051
	(0.100)	(0.091)	(0.091)	(0.097)	(0.074)	(0.070)	(0.070)	(0.075)
rNRSP_7074	0.014	0.088	0.096	0.057	-0.142*	-0.132*	-0.141*	-0.130
	(0.105)	(0.096)	(0.097)	(0.103)	(0.078)	(0.075)	(0.075)	(0.081)
rNRSP_7579	-0.181	-0.101	-0.103	-0.221*	-0.182**	-0.156*	-0.164*	-0.174*
	(0.115)	(0.107)	(0.107)	(0.115)	(0.092)	(0.091)	(0.091)	(0.099)
rage	0.061***	0.059**	0.059**	0.028	0.047***	0.051***	0.051***	0.049***
	(0.022)	(0.024)	(0.024)	(0.025)	(0.017)	(0.017)	(0.017)	(0.019)
rage_sq	-0.093***	-0.093***	-0.092***	-0.067***	-0.077***	-0.082***	-0.083***	-0.081***
	(0.018)	(0.019)	(0.019)	(0.020)	(0.014)	(0.014)	(0.014)	(0.015)
rprimary	0.004	0.017	0.007	0.021	0.023	0.050	0.048	0.063*
	(0.038)	(0.038)	(0.038)	(0.042)	(0.034)	(0.035)	(0.035)	(0.037)
rsecondary	-0.007	0.004	-0.007	0.008	-0.061	-0.011	-0.012	0.010
	(0.044)	(0.045)	(0.045)	(0.049)	(0.042)	(0.043)	(0.043)	(0.047)
rhighabove	-0.033	-0.013	-0.028	-0.021	-0.052	0.025	0.023	0.016
	(0.062)	(0.062)	(0.063)	(0.067)	(0.073)	(0.075)	(0.075)	(0.078)
rmarried	0.416***	0.430***	0.357***	0.310***	0.209***	0.235***	0.272***	0.304***
	(0.047)	(0.048)	(0.105)	(0.114)	(0.035)	(0.036)	(0.079)	(0.086)
hchild	-0.036***	-0.029**	0.073***	0.072***	-0.027***	-0.026**	0.028	0.029
	(0.012)	(0.012)	(0.023)	(0.025)	(0.010)	(0.011)	(0.021)	(0.023)
hhhres	-0.022***	-0.030***	0.014	0.012	-0.010	-0.027***	-0.002	-0.001
	(0.008)	(0.009)	(0.011)	(0.012)	(0.007)	(0.007)	(0.009)	(0.010)
lnhhadurbl	0.028***	0.027***	-0.003	-0.002	-0.002	-0.001	-0.003	-0.002
	(0.007)	(0.007)	(0.008)	(0.008)	(0.005)	(0.005)	(0.006)	(0.007)
y2013	-0.011	-0.046	-0.063	-0.044	0.052	0.046	0.043	0.059*
	(0.045)	(0.042)	(0.043)	(0.046)	(0.033)	(0.032)	(0.032)	(0.035)
y2015	-0.001	-0.042	-0.055	-0.040	0.019	0.001	0.008	0.004
	(0.046)	(0.043)	(0.044)	(0.047)	(0.033)	(0.033)	(0.033)	(0.036)
urban_nbs	-0.731***	-0.723***	-0.743***	-0.623***	-0.884***	-0.877***	-0.882***	-0.657***
	(0.077)	(0.080)	(0.081)	(0.112)	(0.064)	(0.067)	(0.067)	(0.094)
_cons	0.554	0.171	0.048	0.872	0.723	0.245	0.254	0.136
	(0.697)	(0.748)	(0.750)	(0.798)	(0.498)	(0.529)	(0.524)	(0.581)
Provincial Dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$\overline{X_{ij}}$	No	No	Yes	Yes	No	No	Yes	Yes
Community Controls	No	No	No	Yes	No	No	No	Yes
$\rho_{NRSP-work}$	0.077	0.036	0.030	0.064	0.113***	0.101***	0.104***	0.107***
	(0.059)	(0.051)	(0.051)	(0.054)	(0.041)	(0.038)	(0.038)	(0.041)
$\rho_{work-hrs}$	-0.013	-0.047	-0.032	-0.033	-0.217***	-0.160***	-0.151***	-0.144***
	(0.033)	(0.031)	(0.033)	(0.033)	(0.035)	(0.036)	(0.037)	(0.038)
Observations	18636	18635	18635	16018	21227	21227	21227	18330
log likelihood	-86489	-85601	-85569	-73576	-91174	-90021	-90001	-78014

Notes: ibis.

Individuals coming from communities that have older labour forces, lower income levels, more rural workers, and fewer high school degree holders are more likely to be working and report longer hours of work.

As shown in Table 1.9, the covariance between error terms across waves, or $\sigma_h rs$, is significant in the working hour equation, suggesting that the standard errors should be clustered by individual level if estimating a pooled model. The significant and negative $\rho_{NRSP-hrs}$ for men (columns 5 -8 in Table 1.9) suggests that the unobservable factors predicting men's participation in the NRSP are negatively correlated with those predicting more hours of working. In other words, men working for longer hours are less likely to participate in the NRSP than men who work less intensively, given that they are both in the labour force and are similar in the characteristics that we control for. Later heterogeneity analyses using individuals from different job sectors show that the effect comes from the fact that younger non-agricultural workers who work for longer hours are significantly less likely to join the NRSP compared with agricultural workers. Not accounting for the endogeneity of pension participation would bias the estimation of its effect on individual labour supply behaviour. For women, $\rho_{NRSP-hrs}$ is insignificant and of small magnitude, suggesting that women's decision to participate in the NRSP is not related to their decision on hours of working.

For people staying in the labour force, Table 1.8 shows that participation in the NRSP significantly increases the hours of working for current male workers who are below the pension eligible age of 60 and need to contribute to their pension accounts. Heterogeneity analyses show that the effect mainly comes from age-ineligible, non-agricultural workers increasing their hours of working (column 2 in Table 1.12), probably to save for their pensions. Given that the average weekly working hours for age-ineligible male workers are 49.1 hours, they increase their hours by 5.86% to 7.67% based on estimates from column 3. The positive effect on labour intensity will be downwardly biased if not accounting for the endogeneity of pension participation. In contrast, females participants do not significantly change their labour supply intensity.

In summary, the negative effects of pension participation on women's working probability and hours of working, though not very significant, suggest that an increase in pension benefits and expected SSW substitutes for the current labour income and results in female workers reducing their hours of working or completely stopping to work. The negative labour supply responses also differ across age groups, and older contributors tend to respond more than younger contributors do. This can be explained by a life-cycle pattern of the pension effect that our theoretical models predict. Older workers approaching the pension-eligible age have less uncertainty about the future and the survival risk, and need to contribute for less years than the younger participants do. The lack of effect on male pensioners' working probability or working hours can be explained by the low levels of basic benefits that the NRSP offers. It might only see some behavioural changes in the low-income groups and

Table 1.9: Effect of NRSP participation on Hours of Working for the Current Workers

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
rNRSP_below45	4.989**	3.890*	3.982*	3.933	1.652	0.550	0.515	1.921
	(2.405)	(2.284)	(2.289)	(2.428)	(1.873)	(1.777)	(1.774)	(1.892)
rNRSP_4549	4.632***	3.727***	3.764***	4.999***	2.246	0.337	0.374	1.505
	(1.515)	(1.378)	(1.381)	(1.482)	(1.425)	(1.334)	(1.331)	(1.408)
rNRSP_5054	3.707**	2.865**	2.876**	3.882***	1.106	-0.909	-0.871	0.351
	(1.470)	(1.338)	(1.342)	(1.443)	(1.418)	(1.337)	(1.334)	(1.410)
rNRSP_5559	3.718**	2.871**	2.902**	4.057***	1.253	-0.501	-0.428	0.810
	(1.457)	(1.326)	(1.331)	(1.430)	(1.395)	(1.316)	(1.315)	(1.397)
rNRSP_6064	3.011**	1.782	1.813	3.381**	1.530	-0.435	-0.368	1.003
	(1.494)	(1.357)	(1.362)	(1.461)	(1.407)	(1.321)	(1.320)	(1.406)
rNRSP_6569	2.325	1.273	1.298	1.870	1.136	-0.628	-0.692	1.055
	(1.602)	(1.465)	(1.470)	(1.576)	(1.563)	(1.471)	(1.468)	(1.576)
rNRSP_7074	2.383	1.321	1.376	2.237	1.484	-0.275	-0.433	1.111
	(1.847)	(1.730)	(1.732)	(1.864)	(1.801)	(1.737)	(1.736)	(1.853)
rNRSP_7579	0.336	-0.190	-0.240	0.789	0.312	-2.098	-2.272	-0.320
	(2.222)	(2.129)	(2.127)	(2.342)	(2.360)	(2.267)	(2.264)	(2.502)
rage	1.178***	1.092***	1.112***	1.074**	0.185	0.307	0.316	0.528*
	(0.373)	(0.368)	(0.370)	(0.425)	(0.295)	(0.282)	(0.282)	(0.294)
rage_sq	-1.426***	-1.330***	-1.360***	-1.323***	-0.562**	-0.691***	-0.705***	-0.895***
	(0.320)	(0.315)	(0.318)	(0.363)	(0.258)	(0.248)	(0.249)	(0.260)
rprimary	-0.410	-0.600	-0.539	-0.543	0.026	-0.247	-0.302	-0.336
	(0.583)	(0.573)	(0.575)	(0.621)	(0.643)	(0.641)	(0.642)	(0.686)
rsecondary	-0.457	-0.965	-0.894	-0.926	0.066	0.507	0.516	1.213
	(0.646)	(0.646)	(0.647)	(0.696)	(0.804)	(0.801)	(0.802)	(0.858)
rmarried	5.408***	4.671***	0.133	0.105	1.576*	1.835**	-1.138	-0.985
	(0.916)	(0.908)	(2.440)	(2.596)	(0.847)	(0.840)	(2.099)	(2.328)
hchild	-0.353*	-0.281	-0.450	-0.326	-0.038	-0.036	-0.117	-0.199
	(0.207)	(0.210)	(0.515)	(0.561)	(0.216)	(0.220)	(0.535)	(0.573)
hhhres	-0.184	0.056	-0.142	-0.167	-0.641***	-0.601***	0.037	0.027
	(0.139)	(0.140)	(0.224)	(0.243)	(0.138)	(0.140)	(0.212)	(0.227)
lnhhadurbl	0.181	0.213*	0.525***	0.577***	0.140	0.150	0.153	0.190
	(0.122)	(0.121)	(0.161)	(0.174)	(0.127)	(0.126)	(0.167)	(0.178)
y2013	-1.694**	-1.348**	-1.376**	-1.800**	-1.181	-0.393	-0.432	-1.179
	(0.720)	(0.676)	(0.677)	(0.744)	(0.723)	(0.694)	(0.694)	(0.752)
y2015	-1.831**	-1.282*	-1.385**	-1.563**	-1.318*	-0.455	-0.169	-0.469
	(0.717)	(0.672)	(0.680)	(0.741)	(0.725)	(0.699)	(0.709)	(0.761)
urban_nbs	6.411***	5.527***	5.539***	3.610*	6.892***	5.240***	5.166***	4.860**
	(1.388)	(1.432)	(1.438)	(1.958)	(1.782)	(1.810)	(1.814)	(2.306)
_cons	20.777*	29.945***	30.249***	32.752**	48.882***	58.611***	58.532***	53.052***
	(10.827)	(11.032)	(11.084)	(12.790)	(8.403)	(8.506)	(8.509)	(9.040)
Provincial Dummies	No	Yes	Yes	Yes	No	Yes	Yes	Yes
$\overline{X_{ij}}$	No	No	Yes	Yes	No	No	Yes	Yes
Community Controls	No	No	No	Yes	No	No	No	Yes
$ln\sigma_{hrs}$	3.190***	3.178***	3.178***	3.178***	3.249***	3.230***	3.229***	3.227***
	(0.007)	(0.007)	(0.007)	(0.008)	(0.006)	(0.006)	(0.006)	(0.007)
$\rho_{NRSP-hrs}$	-0.088**	-0.068**	-0.068**	-0.094***	-0.054*	0.006	0.006	-0.023
	(0.035)	(0.031)	(0.032)	(0.034)	(0.030)	(0.028)	(0.028)	(0.030)
$\rho_{work-hrs}$	-0.013	-0.047	-0.032	-0.033	-0.217***	-0.160***	-0.151***	-0.144***
	(0.033)	(0.031)	(0.033)	(0.033)	(0.035)	(0.036)	(0.037)	(0.038)
Observations	18636	18635	18635	16018	21227	21227	21227	18330
log likelihood	-86489	-85601	-85569	-73576	-91174	-90021	-90001	-78014

Notes: ibis.

workers in poor health status or suffering from high opportunity of working.

1.7.3 Labour Supply Decision and NRSP Contribution Decision

In this section, we study the relationship between individual contribution decision and individual labour supply decisions by estimating the effect of moving from contributing at the minimum requirement level (referred to as low contribution) to contributing above the minimum requirement level (referred to as high contribution) on the NRSP participants' labour supply outcomes. The key variable of interest is an indicator of high contribution that equals 1 for high contribution and 0 for low contribution, missing for non-participants. Duration of the NRSP programs in local communities are used to predict the probability of making high contribution, under the assumption that the longer the programme has been in place, the more likely people will participate and will contribute at a higher level to their pensions after a build-up of knowledge about the programme and trust in the local governments.

We estimate jointly in a trivariate model the three equations of NRSP participation decision, NRSP contribution decision, and one of the two labour supply decisions of whether working or not and hours of work conditional on working. Estimation is separated for men and women, and the sample is restricted to rural residents aged below 60, not yet formally retired and not missing information about pension participation status and labour supply status. Due to the small sample size of high contributors, we do not study the heterogeneous effects across age groups.

Table 1.10 reports the estimation outputs for the four models. Standard errors are clustered in individual levels, and we also control for provincial dummies and individual-level means of the time-varying exogenous covariates. Because the estimations of the NRSP participation equation ($NRSP_{ijt}$) and the NRSP contribution equation (HC_{ijt}) use the same sample and produce similar outputs in trivariate model that predicts working probability and in trivariate model that predicts hours of working, we only report their estimated coefficients in the trivariate model that predicts working probability. However, the correlations between unobservables are different and we report the correlation between unobservables predicting pension related outcomes and unobservables that predict hours of work in columns 4 and 8. Columns 2 and 6 in Table 1.10 report estimation of the contribution equation (equation 1.9). Community-level duration of the NRSP significantly predicts the annual contribution levels of participants. Recall that $NRSP \leq 0y$ equals 1 for communities that had not started the NRSP in 2011 or had just started it earlier that year, and equals 0 for the rest of communities in 2011 or for all communities in 2013 or 2015, when the NRSP had reached nation-wide coverage. $NRSP(0, 2y]$ equals 1 for communities that had started the NRSP and had it in place for 2 years or less

Table 1.10: Estimation for the NRSP contribution levels

	Men				Women			
	$NRSP_{ijt}$	HC_{ijt}	Work	Hours	$NRSP_{ijt}$	HC_{ijt}	Work	Hours
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
NRSP in place	1.886***				1.884***			
	(0.064)				(0.060)			
NRSP$\leq$ 0y		-2.901***				-1.210***		
		(0.163)				(0.323)		
NRSP (0,2y]		-0.908***				-1.131***		
		(0.098)				(0.120)		
NRSP (2,4y]		-0.284***				-0.346***		
		(0.059)				(0.069)		
HC_{ijt}			-0.298	-3.291			-0.151	-6.854**
			(0.284)	(3.460)			(0.252)	(3.268)
rage	-0.012	0.045	-0.161	1.388	0.019	-0.048	-0.052	-0.357
	(0.044)	(0.093)	(0.182)	(1.851)	(0.029)	(0.060)	(0.074)	(1.127)
rage_sq	0.034	-0.021	0.120	-1.715	0.008	0.052	0.006	-0.102
	(0.043)	(0.089)	(0.173)	(1.783)	(0.029)	(0.060)	(0.073)	(1.117)
rprimary	0.043	0.094*	0.096	0.320	0.041	0.114**	0.211***	1.149
	(0.040)	(0.056)	(0.087)	(1.056)	(0.034)	(0.051)	(0.059)	(0.976)
rsecondary	0.078**	0.120**	0.142*	0.489	0.077**	0.140**	0.113*	0.998
	(0.040)	(0.053)	(0.085)	(1.026)	(0.037)	(0.055)	(0.061)	(1.065)
rhighabove	0.027	0.101	0.072	0.213	0.046	0.069	0.155	1.222
	(0.050)	(0.065)	(0.106)	(1.284)	(0.063)	(0.094)	(0.100)	(1.824)
rmarried	0.405*	0.283	0.301	-1.499	0.090	-0.178	0.500*	0.713
	(0.237)	(0.327)	(0.441)	(6.160)	(0.174)	(0.283)	(0.278)	(5.338)
hchild	0.014	0.029	-0.142	0.188	0.071*	-0.071	-0.046	0.018
	(0.045)	(0.062)	(0.096)	(1.212)	(0.041)	(0.069)	(0.069)	(1.321)
hhhres	-0.024	0.037	0.030	-0.559	-0.020	0.006	0.004	0.318
	(0.018)	(0.023)	(0.036)	(0.450)	(0.015)	(0.023)	(0.024)	(0.446)
lnhhadurbl	0.019	0.035*	-0.021	-0.080	0.018	0.004	0.032	0.093
	(0.014)	(0.020)	(0.031)	(0.388)	(0.012)	(0.020)	(0.020)	(0.384)
y2013	0.376***	-0.397***	-0.175*	-0.886	0.413***	-0.855***	-0.128	-1.245
	(0.040)	(0.067)	(0.101)	(1.202)	(0.037)	(0.069)	(0.079)	(1.317)
y2015	0.298***	-0.526***	-0.074	-0.684	0.335***	-1.045***	-0.138*	-0.082
	(0.041)	(0.099)	(0.105)	(1.215)	(0.038)	(0.107)	(0.074)	(1.273)
urban_nbs	-0.209***	0.203**	-0.744***	7.892***	-0.176**	0.300***	-0.743***	6.633**
	(0.077)	(0.099)	(0.155)	(2.371)	(0.071)	(0.109)	(0.121)	(2.610)
_cons	-2.225**	-2.200	5.633	30.881	-3.048***	1.881	3.455*	82.430***
	(1.134)	(2.403)	(4.761)	(47.979)	(0.730)	(1.548)	(1.924)	(28.906)
σ_{hrs}				3.154***				3.234***
				(0.012)				(0.012)
$\rho_{NRSP-HC}$	1.670***	1.670***	1.670***	1.592***	-0.542***	-0.542***	-0.542***	-0.496***
	(0.348)	(0.348)	(0.348)	(0.284)	(0.156)	(0.156)	(0.156)	(0.152)
$\rho_{NRSP-work/hrs}$	0.032	0.032	0.032	-0.042	-0.159	-0.159	-0.159	0.122
	(0.146)	(0.146)	(0.146)	(0.074)	(0.136)	(0.136)	(0.136)	(0.087)
$\rho_{HC-work/hrs}$	0.194	0.194	0.194	0.078	0.127	0.127	0.127	0.156*
	(0.147)	(0.147)	(0.147)	(0.073)	(0.152)	(0.152)	(0.152)	(0.081)
Observations	10314	10314	10314	10314	12078	12078	12078	12078
log likelihood	-8677	-8677	-8677	-28201	-11187	-11187	-11187	-31035

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual levels. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to age-ineligible rural residents not formally retired and not missing in the outcome variables.

in the survey year. $NRSP(2,4y]$ equals 1 for communities that had had the NRSP in place for more than 2 years but less than or equal to 4 years in the survey year. $NRSP > 4y$ or communities that had had the NRSP in place for more than 4 years in the survey year are used to construct the reference group.

Coefficients of the community-level duration of the NRSP suggest that participants coming from communities that started the NRSP later, or have a shorter duration of the NRSP, are significantly less likely to contribute above the minimum requirement to their pension accounts. Participants who have completed primary school or secondary school are more likely to contribute above the minimum requirement than participants who have not finished primary school. Participants living in urban areas also contribute more.

The significant and negative correlation between unobserved effects that predict NRSP participation and high levels of contribution for women (columns 5-8 in Table 1.10), or $\rho_{NRSP-HC}$, suggests that age-ineligible women who participate in the NRSP tend to contribute at the minimum requirement. In contrast, $\rho_{NRSP-HC}$ is significant and positive for men, suggesting that age-ineligible male participants tend to contribute above the minimum requirement.

$\rho_{NRSP-work}$ is still positive and of similar size for age-ineligible men (columns 1-3 in Table 1.10), compared to its estimate for the whole sample of men (column 3 in Table 1.7). Although it is non-significant for both models, it indicates that age-ineligible men are not less likely to participate than the older group. As for women, the sign of $\rho_{NRSP-work}$ reverses when we compare the model using all age groups (column 7 in Table 1.7) with that using only the age-ineligible group (columns 5-7 in Table 1.10). It suggests that among age-ineligible women, those not working are more likely to participate in the NRSP (though not significant), while among the age-eligible women, those still working are more likely to participate.

The correlations between unobserved effects that predict NRSP participation and hours of work, or $\rho_{NRSP-hrs}$ in columns 4 and 8, show similar signs as what we estimate using the whole sample of men or women (columns 3 and 7 in Table 1.7). Although being insignificant, the negative $\rho_{NRSP-hrs}$ for men (column 4) suggests that age-ineligible men working for longer hours are less likely to participate in the NRSP. Heterogeneity analyses in the following section among individuals from different job sectors show that the effect comes younger non-agricultural workers who work for longer hours and are significantly less likely to join the NRSP compared with younger agricultural workers. $\rho_{NRSP-hrs}$ is positive for age-ineligible women (column 8), suggesting that women who work for longer hours are more likely to participate in the NRSP than women who work less intensively, given that they are both in the labour force and are similar in other observable characteristics. Heterogeneity analyses show that the effect comes from age-ineligible women doing agricultural work. Not accounting for the omitted variable bias would underestimate the negative labour supply adjustment to high contribution levels.

$rho_{HC-work}$ is positive but not significant for either men or women, suggesting that high contributors are not significantly more likely to participate in the labour force compared to the low contributors. Among those in the labour force, the positive rho_{HC-hrs} shows that workers reporting more hours of labour supply are more likely to contribute at a higher level to their pensions, especially for female workers. Not accounting for the endogeneity of contribution levels will downwardly bias its negative effect on contributors' labour supply.

Estimates of the coefficients show that making a high contribution does not significantly affect the individual probability of labour participation, but reduces hours of working for the current female workers. Given that the average weekly working hours for the sampled female workers aged below 60 are 42.92 hours, the effect amounts to a 15.97% decline in labour supply intensity. Recall that the trivariate models in the previous section also find a negative effect of pension participation on age-ineligible women's labour participation rate and hours of working, though not very significant. Table 1.10 suggests that female participants who contribute at a higher level to their pensions withdraw further from the formal labour supply. For men, high contributors reduce their hours of working and are less likely to work, though the effects are not significant. Models in the previous section find that participation in the NRSP significantly increases the hours of working for men, especially for non-agricultural workers. Results in Table 1.10 suggest that the positive effects on working hours only apply for low contributors.

1.8 Heterogeneity Analysis

In this section, we separate the total sample into demographic subgroups based on job sectors, education levels, quantiles of household durable assets' values, and re-estimate the trivariate model of the probability of participating in the NRSP, probability of working and hours of work for each subgroups. The aim is to explore the heterogeneous effects of NRSP participation on labour supply behaviours across population subgroups, and to better understand the mechanism behind the effects.

1.8.1 Labour Supply Transitions and NRSP Participation Decision

Based on the characteristics of their main jobs, we categorize current workers as either agricultural workers or non-agricultural workers, with the latter group including the self-employed and individuals working in non-agricultural firms. Instead of studying the single transition between working and non-working in the previous section, we look at three types of transitions including the transition between agricultural work and non-working (columns 1 and 5 in Table 1.11), the transition between non-agricultural work and non-working (columns 2 and 6 in Table 1.11), the transition between agricultural work and non-agricultural work (columns 3 7 in Table 1.11). We do not estimate a multinomial probit equation of labour market transitions in the trivariate framework due to computational limit. Columns 4 and 8 in Tables 1.11 and 1.12 are estimates from extracted from Tables 1.8 and 1.9 that use the whole sample and pools agricultural and non-agricultural workers together.

Hours of working are still missing for non-workers. The sample is restricted to agricultural workers and non-workers in columns 1 and 5 when we study the transition between farming work and non-working (Table 1.11) and farming working hours (Table 1.12), to non-agricultural workers and non-workers in columns 2 and 6 when we study the transition between non-farming work and non-working (Table 1.11) and non-farming working hours (Table 1.12), and to all current workers including agricultural and non-agricultural workers in columns 3 and 7 when we study the transition between farming and non-farming work (Table 1.11) and the average hours of working (Table 1.12).

The estimated probability of NRSP participation for different subgroups are reported in Table A.1 in the Appendix A. Demographic variables show similar effects compared to how they are in models using the whole sample (Table 1.7). Gender difference is small in terms of the significant levels and sizes of the covariates that determine the probability of participating in the NRSP. Non-agricultural workers are less likely to join the NRSP compared to agricultural workers, as suggested by the lower magnitude of the effect of the placement of the NRSP in models that exclude agricultural workers.

The positive and significant $\rho_{NRSP-work}$ in columns 3 and 7 in Table 1.11 suggests that agricultural workers are more likely to participate than non-agricultural workers do, though it is only significant for men. The positive and significant $\rho_{NRSP-work}$ in columns 1 and 5 in Table 1.11 suggests that agricultural workers are more likely to participate than non-workers do, and the effect is more substantial and significant among women. The negative $\rho_{NRSP-work}$ in columns 2 and 6 indicates that non-agricultural workers are less likely to join the NRSP than non-workers do, especially for men. In summary, agricultural workers are most likely and non-agricultural workers are least likely to participate in the NRSP among all eligible men. Among male non-agricultural workers, NRSP participants tend to be those who work for significantly less hours than the others, as suggested by the negative and significant $\rho_{NRSP-hrs}$ in the model that estimates non-agricultural working hours (column 2 in Table 1.12). The significant and negative $\rho_{work-hrs}$ in columns 3 and 7 suggests that for both men and women, farmers work for less hours than non-farmers do. The negative $\rho_{work-hrs}$ in models of agricultural or non-agricultural working hours shows that older workers

Table 1.11: Labour Market Transitions of NRSP Participants by Job Sectors

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	F-R	NF-Ret	F-NF	W-R	F-R	NF-Ret	F-NF	W-R
rNRSP_below45	0.087	0.156	-0.057	-0.003	0.245*	0.432***	-0.097	0.118
	(0.274)	(0.276)	(0.137)	(0.223)	(0.131)	(0.167)	(0.117)	(0.117)
rNRSP_4549	0.166	0.231	-0.128	0.113	0.047	0.250**	-0.049	-0.009
	(0.131)	(0.152)	(0.088)	(0.107)	(0.078)	(0.126)	(0.094)	(0.074)
rNRSP_5054	0.154	0.234	-0.109	0.114	-0.030	0.011	0.061	-0.122*
	(0.117)	(0.143)	(0.087)	(0.098)	(0.075)	(0.123)	(0.096)	(0.071)
rNRSP_5559	0.139	0.264*	-0.119	0.115	-0.044	-0.071	0.096	-0.102
	(0.107)	(0.143)	(0.086)	(0.094)	(0.070)	(0.124)	(0.098)	(0.068)
rNRSP_6064	0.127	0.214	-0.081	0.091	0.027	-0.077	0.057	-0.028
	(0.105)	(0.142)	(0.089)	(0.091)	(0.069)	(0.125)	(0.102)	(0.067)
rNRSP_6569	0.054	0.199	-0.197**	0.061	-0.008	-0.001	-0.053	-0.029
	(0.104)	(0.147)	(0.100)	(0.091)	(0.072)	(0.133)	(0.114)	(0.070)
rNRSP_7074	0.097	0.257	-0.314***	0.096	-0.123	-0.029	-0.160	-0.141*
	(0.115)	(0.157)	(0.120)	(0.097)	(0.079)	(0.143)	(0.138)	(0.075)
rNRSP_7579	-0.136	0.325*	-0.574***	-0.103	-0.138	-0.084	-0.091	-0.164*
	(0.126)	(0.192)	(0.171)	(0.107)	(0.095)	(0.186)	(0.201)	(0.091)
rNRSP_above80	0.166	0.623**	-0.646**	0.169	0.052	0.258	-0.379	0.003
	(0.161)	(0.285)	(0.283)	(0.136)	(0.129)	(0.231)	(0.269)	(0.123)
rage	0.094***	0.075	-0.009	0.059**	0.089***	-0.041	0.095***	0.051***
	(0.025)	(0.046)	(0.029)	(0.024)	(0.017)	(0.034)	(0.025)	(0.017)
rage_sq	-0.112***	-0.135***	0.046*	-0.092***	-0.109***	-0.020	-0.057***	-0.083***
	(0.020)	(0.038)	(0.026)	(0.019)	(0.014)	(0.029)	(0.022)	(0.014)
rprimary	-0.017	0.049	-0.092**	0.007	0.028	0.187***	-0.193***	0.048
	(0.042)	(0.054)	(0.037)	(0.038)	(0.037)	(0.049)	(0.039)	(0.035)
rsecondary	-0.065	0.098*	-0.167***	-0.007	-0.080*	0.192***	-0.323***	-0.012
	(0.053)	(0.059)	(0.039)	(0.045)	(0.047)	(0.059)	(0.047)	(0.043)
rhighabove	-0.143*	0.157**	-0.357***	-0.028	-0.114	0.388***	-0.487***	0.023
	(0.076)	(0.079)	(0.053)	(0.063)	(0.087)	(0.097)	(0.084)	(0.075)
rmarried	0.359***	0.414**	0.034	0.357***	0.288***	0.043	0.169	0.272***
	(0.110)	(0.179)	(0.151)	(0.105)	(0.085)	(0.143)	(0.125)	(0.079)
hchild	0.071***	0.100***	-0.035	0.073***	0.023	0.010	-0.019	0.028
	(0.023)	(0.038)	(0.025)	(0.023)	(0.022)	(0.036)	(0.033)	(0.021)
hhhres	0.012	0.017	-0.013	0.014	-0.001	-0.011	0.006	-0.002
	(0.012)	(0.017)	(0.011)	(0.011)	(0.010)	(0.016)	(0.013)	(0.009)
lnhhadurbl	-0.011	0.009	-0.017*	-0.003	-0.002	-0.002	-0.003	-0.003
	(0.008)	(0.013)	(0.009)	(0.008)	(0.007)	(0.011)	(0.010)	(0.006)
y2013	-0.120**	-0.074	-0.050	-0.063	0.003	0.042	-0.073	0.043
	(0.051)	(0.061)	(0.040)	(0.043)	(0.034)	(0.054)	(0.045)	(0.032)
y2015	-0.173***	0.062	-0.278***	-0.055	-0.084**	0.159***	-0.296***	0.008
	(0.056)	(0.062)	(0.042)	(0.044)	(0.036)	(0.053)	(0.045)	(0.033)
urban_nbs	-1.458***	-0.258***	-1.289***	-0.743***	-1.448***	-0.058	-1.435***	-0.882***
	(0.170)	(0.100)	(0.111)	(0.081)	(0.098)	(0.082)	(0.108)	(0.067)
_cons	-1.333*	-0.709	0.789	0.048	-1.184**	2.311**	-2.055***	0.254
	(0.788)	(1.388)	(0.851)	(0.750)	(0.525)	(1.020)	(0.742)	(0.524)
$\rho_{NRSP-work}$	0.037	-0.064	0.132***	0.030	0.094**	-0.003	0.072	0.104***
	(0.058)	(0.082)	(0.051)	(0.051)	(0.039)	(0.070)	(0.055)	(0.038)
$\rho_{work-hrs}$	-0.301	-0.033*	-0.347***	-0.032	-0.164***	-0.060*	-0.390***	-0.151***
	(0.203)	(0.020)	(0.013)	(0.033)	(0.052)	(0.035)	(0.015)	(0.037)
Observations	12825	8633	15385	18635	17675	9009	15070	21227
Log likelihood	-56381	-32571	-84953	-85569	-71793	-23453	-81999	-90001

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, belonging to the specific labour groups studied in each model, and not missing in the outcome variables.

staying in the labour force tend to reduce their hours of working, especially for older agricultural workers and female workers.

Estimates of the NRSP participation effects show significant gender differences. Male participants aged above 65 are significantly more likely to move from agricultural jobs to non-agricultural jobs. Women tend to stop working completely after joining the NRSP, but the effects are not significant. For those staying in the labour force, male non-agricultural workers aged below 60 significantly increase their working hours after joining the NRSP. Given that the average working hours for age-ineligible male workers doing non-agricultural jobs is 55.40 hours per week, the positive NRSP effect accounts for an 8.37% to 14.13% increase. Female agricultural workers aged above 50 reduce their working hours, as shown in column 5, though the decline is not significant. Female workers start to reduce working hours after 50 despite the fact that they still need to contribute to their pension accounts. It suggests that the expectation of higher future pension income outweighs the current income effect in determining the labour supply responses of the age-ineligible female workers.

Estimates of control variables in Table 1.11 suggest that older, lower-educated individuals living in rural areas are more likely to work in agricultural sectors. There is a time trend of transiting from agricultural to non-agricultural sectors for both men and women. Table 1.12 shows that agricultural workers attaining high education work for fewer hours, compared to the reference group of individuals who have not completed primary education. Values of household durable assets are positively related to hours of agricultural work but not hours of non-agricultural work. There is a time trend of declining hours of working for non-agricultural workers. Non-agricultural workers living in urban areas work for longer hours than non-agricultural workers living in rural areas do.

Table 1.12: Effects of NRSP Participation on the Working Hours by Job Sectors

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	farm	nonfarm	all	all	farm	nonfarm	all	all
rNRSP_below45	0.637	7.829**	4.015*	3.982*	1.150	1.855	0.452	0.515
	(3.472)	(3.164)	(2.282)	(2.289)	(2.065)	(3.466)	(1.769)	(1.774)
rNRSP_4549	2.968	4.817**	3.751***	3.764***	0.137	3.558	0.109	0.374
	(2.551)	(2.148)	(1.371)	(1.381)	(1.523)	(3.033)	(1.335)	(1.331)
rNRSP_5054	1.331	4.639**	2.876**	2.876**	-0.419	2.155	-1.262	-0.871
	(2.362)	(2.177)	(1.330)	(1.342)	(1.516)	(3.174)	(1.337)	(1.334)
rNRSP_5559	0.461	5.850***	2.914**	2.902**	-0.495	3.922	-0.706	-0.428
	(2.356)	(2.241)	(1.321)	(1.331)	(1.484)	(3.358)	(1.317)	(1.315)
rNRSP_6064	0.453	2.895	1.840	1.813	-0.286	0.295	-0.497	-0.368
	(2.299)	(2.537)	(1.351)	(1.362)	(1.484)	(3.558)	(1.322)	(1.320)
rNRSP_6569	-1.343	4.494	1.296	1.298	-1.371	1.547	-0.821	-0.692
	(2.390)	(3.267)	(1.461)	(1.470)	(1.611)	(4.457)	(1.468)	(1.468)
rNRSP_7074	-0.344	-2.106	1.415	1.376	0.358	-14.214**	-0.810	-0.433
	(2.482)	(5.991)	(1.719)	(1.732)	(1.867)	(5.826)	(1.735)	(1.736)
rNRSP_7579	-1.992	-5.865	-0.302	-0.240	-3.488	-3.172	-3.100	-2.272
	(2.850)	(7.526)	(2.123)	(2.127)	(2.373)	(7.327)	(2.256)	(2.264)
rNRSP_above80	4.761	17.686	9.571***	9.527***	3.008	0.225	3.035	3.793
	(4.039)	(17.290)	(3.613)	(3.615)	(3.540)	(11.581)	(3.540)	(3.541)
rage	0.716	0.495	1.150***	1.112***	0.947***	0.759	0.485*	0.316
	(0.585)	(0.798)	(0.364)	(0.370)	(0.352)	(0.607)	(0.282)	(0.282)
rage_sq	-0.832	-0.473	-1.403***	-1.360***	-1.103***	-1.010*	-0.916***	-0.705***
	(0.522)	(0.747)	(0.310)	(0.318)	(0.307)	(0.578)	(0.248)	(0.249)
rprimary	-1.403*	0.460	-0.495	-0.539	-2.064***	1.506	-0.187	-0.302
	(0.717)	(0.998)	(0.575)	(0.575)	(0.687)	(1.354)	(0.642)	(0.642)
rsecondary	-1.997**	-0.157	-0.849	-0.894	-1.364	0.765	0.482	0.516
	(0.819)	(1.053)	(0.647)	(0.647)	(0.866)	(1.528)	(0.804)	(0.802)
rhighabove	-3.173**	-2.229*	-1.634*	-1.666*	0.167	0.479	2.054	2.003
	(1.291)	(1.331)	(0.941)	(0.940)	(1.765)	(2.156)	(1.399)	(1.399)
rmarried	-0.754	1.622	0.174	0.133	1.931	-15.347**	-0.587	-1.138
	(2.820)	(4.804)	(2.435)	(2.440)	(2.112)	(6.356)	(2.085)	(2.099)
hchild	-0.621	-1.016	-0.425	-0.450	-0.275	1.070	-0.107	-0.117
	(0.621)	(1.084)	(0.511)	(0.515)	(0.549)	(1.436)	(0.528)	(0.535)
hhhres	-0.386	0.144	-0.147	-0.142	0.054	-0.118	0.022	0.037
	(0.286)	(0.368)	(0.224)	(0.224)	(0.218)	(0.594)	(0.211)	(0.212)
lnhhadurbl	0.712***	-0.071	0.519***	0.525***	0.170	-0.169	0.155	0.153
	(0.195)	(0.307)	(0.161)	(0.161)	(0.173)	(0.438)	(0.166)	(0.167)
y2013	-0.374	-2.945***	-1.373**	-1.376**	-0.365	-2.509*	-0.304	-0.432
	(1.319)	(0.998)	(0.674)	(0.677)	(0.798)	(1.414)	(0.694)	(0.694)
y2015	-1.474	-3.678***	-1.477**	-1.385**	-1.770**	-1.640	-0.189	-0.169
	(1.337)	(0.984)	(0.681)	(0.680)	(0.817)	(1.369)	(0.711)	(0.709)
urban_nbs	-5.866	3.078**	5.424***	5.539***	-9.887***	-2.005	3.506**	5.166***
	(5.884)	(1.378)	(1.431)	(1.438)	(2.868)	(2.015)	(1.786)	(1.814)
_cons	43.483**	41.178*	28.862***	30.249***	33.874***	62.083***	53.740***	58.532***
	(18.471)	(21.705)	(10.921)	(11.084)	(10.845)	(17.953)	(8.538)	(8.509)
$ln\sigma_{hrs}$	3.182***	3.098***	3.178***	3.178***	3.178***	3.225***	3.225***	3.229***
	(0.025)	(0.016)	(0.007)	(0.007)	(0.007)	(0.017)	(0.006)	(0.006)
$\rho_{NRSP-hrs}$	-0.020	-0.120**	-0.068**	-0.068**	0.034	-0.061	0.016	0.006
	(0.056)	(0.058)	(0.031)	(0.032)	(0.035)	(0.068)	(0.029)	(0.028)
$\rho_{work-hrs}$	-0.301	-0.033*	-0.347***	-0.032	-0.164***	-0.060*	-0.390***	-0.151***
	(0.203)	(0.020)	(0.013)	(0.033)	(0.052)	(0.035)	(0.015)	(0.037)
Observations	12825	8633	15385	18635	17675	9009	15070	21227
Log likelihood	-56381	-32571	-84953	-85569	-71793	-23453	-81999	-90001

ibid.

1.8.2 Educational Level

Apart from job sectors, educational level is another factor that determines older workers' retirement ages and labour intensity. Highly educated individuals tend to work in formal sectors, take less physically-demanding jobs and formally retire at younger ages due to the higher levels of private savings and pension incomes.

In this section, we investigate the heterogeneous responses from participants of different educational backgrounds. We categorize individuals into three groups based on the highest educational levels they achieve. They include illiterate individuals not receiving any formal education (columns 1 and 4), individuals not completing or just completing primary school education (columns 2 and 5), individuals completing secondary school or higher levels of education (columns 3 and 6). We re-estimate the trivariate model for each of the subgroups, and separately for men and women. Table 1.13 reports estimates of the effect of participating in the NRSP on individual working probability and Table 1.14 the effect on hours of work. The estimated factors that predict probabilities of participating in the NRSP for different subgroups are reported in Table A.2 in the Appendix A.

The significant and positive $\rho_{NRSP-work}$ for illiterate women, who make up almost half of the sampled women, suggests that among them, current workers are more likely to join the NRSP than non-workers. The significant and negative $\rho_{work-hours}$ suggests that illiterate female workers tend to work less hours if they continue with their work. The negative and significant $\rho_{NRSP-hours}$ for male workers achieving secondary or higher level of education suggests that the NRSP participants among them are more likely to work less hours.

Table 1.12 shows that participating in the NRSP does not significantly change the working probability for men across all educational levels, but does make illiterate women more likely to stop working, especially if they are aged above 70 and eligible to receive the pensions. Table 1.14 shows that among the current workers, age-ineligible men having completed secondary or above levels of education are more likely to work longer hours if they participate in the NRSP compared to those not participating.

Table 1.13: Effects of NRSP Participation on Working Probability by Educational Groups

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	illiterate	primary	secondary	all	illiterate	primary	secondary	all
rNRSP_below45	-0.803*	0.304	-0.168	-0.003	-0.283	0.299	0.622**	0.118
	(0.482)	(0.297)	(0.526)	(0.223)	(0.176)	(0.207)	(0.307)	(0.117)
rNRSP_4549	-0.021	0.144	0.121	0.113	-0.112	0.026	0.181	-0.009
	(0.337)	(0.151)	(0.195)	(0.107)	(0.128)	(0.115)	(0.174)	(0.074)
rNRSP_5054	-0.048	0.148	0.088	0.114	-0.209*	-0.159	0.193	-0.122*
	(0.242)	(0.145)	(0.179)	(0.098)	(0.118)	(0.116)	(0.168)	(0.071)
rNRSP_5559	0.049	0.223*	0.050	0.115	-0.197*	-0.075	0.311*	-0.102
	(0.223)	(0.129)	(0.183)	(0.094)	(0.102)	(0.113)	(0.174)	(0.068)
rNRSP_6064	0.227	0.119	0.024	0.091	-0.128	0.058	0.204	-0.028
	(0.191)	(0.124)	(0.189)	(0.091)	(0.100)	(0.109)	(0.201)	(0.067)
rNRSP_6569	0.091	0.113	-0.077	0.061	-0.163	0.076	0.034	-0.029
	(0.197)	(0.123)	(0.197)	(0.091)	(0.103)	(0.115)	(0.225)	(0.070)
rNRSP_7074	0.056	0.196	-0.151	0.096	-0.192*	-0.167	-0.235	-0.141*
	(0.189)	(0.132)	(0.226)	(0.097)	(0.107)	(0.126)	(0.272)	(0.075)
rNRSP_7579	-0.143	-0.025	-0.375	-0.103	-0.230**	-0.238	0.452	-0.164*
	(0.195)	(0.150)	(0.309)	(0.107)	(0.116)	(0.197)	(0.543)	(0.091)
rNRSP_above80	-0.125	0.438**	0.142	0.169	0.005	-0.215		0.003
	(0.220)	(0.193)	(0.486)	(0.136)	(0.143)	(0.366)		(0.123)
rage	0.047	0.067**	-0.006	0.059**	0.094***	-0.022	0.075	0.051***
	(0.047)	(0.031)	(0.063)	(0.024)	(0.023)	(0.028)	(0.049)	(0.017)
rage_sq	-0.087**	-0.097***	-0.037	-0.092***	-0.115***	-0.024	-0.114**	-0.083***
	(0.036)	(0.025)	(0.053)	(0.019)	(0.019)	(0.024)	(0.045)	(0.014)
rmarried	0.107	0.396***	0.602***	0.357***	0.330***	0.199	0.277	0.272***
	(0.251)	(0.135)	(0.232)	(0.105)	(0.101)	(0.152)	(0.258)	(0.079)
hchild	0.173***	0.037	0.073*	0.073***	0.019	0.060	0.003	0.028
	(0.056)	(0.031)	(0.043)	(0.023)	(0.027)	(0.036)	(0.073)	(0.021)
hhhres	-0.007	0.028*	0.007	0.014	-0.007	-0.002	0.012	-0.002
	(0.025)	(0.016)	(0.021)	(0.011)	(0.013)	(0.016)	(0.024)	(0.009)
lnhhadurbl	0.013	-0.007	-0.011	-0.003	-0.007	-0.001	0.016	-0.003
	(0.016)	(0.010)	(0.018)	(0.008)	(0.008)	(0.012)	(0.020)	(0.006)
y2013	-0.130	-0.083	-0.011	-0.063	0.111**	-0.004	-0.047	0.043
	(0.095)	(0.062)	(0.078)	(0.043)	(0.049)	(0.052)	(0.083)	(0.032)
y2015	-0.143	-0.040	-0.030	-0.055	0.039	0.023	-0.164*	0.008
	(0.096)	(0.063)	(0.084)	(0.044)	(0.051)	(0.053)	(0.088)	(0.033)
urban_nbs	-2.001***	-0.705***	-0.683***	-0.743***	-0.931***	-0.910***	-0.852***	-0.882***
	(0.462)	(0.111)	(0.124)	(0.081)	(0.128)	(0.107)	(0.128)	(0.067)
_cons	0.798	-0.368	2.079	0.048	-1.244	2.485***	-0.104	0.254
	(1.606)	(0.986)	(1.906)	(0.750)	(0.767)	(0.831)	(1.321)	(0.524)
$\rho_{NRSP-work}$	0.131	-0.040	0.081	0.030	0.140**	0.091	-0.007	0.104***
	(0.100)	(0.071)	(0.107)	(0.051)	(0.057)	(0.063)	(0.094)	(0.038)
$\rho_{work-hrs}$	-0.197	0.046	-0.097*	-0.032	-0.199***	-0.117**	-0.021	-0.151***
	(0.188)	(0.043)	(0.053)	(0.033)	(0.073)	(0.047)	(0.126)	(0.037)
Observations	2633	9652	6350	18635	9548	8414	3261	21227
Log likelihood	-10697	-44193	-30466	-85569	-37991	-36904	-14805	-90001

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, belonging to the specific educational groups studied in each model, and not missing in the outcome variables.

Table 1.14: Effects of NRSP Participation on the Working Hours by Educational Groups

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	illiterate	primary	secondary	all	illiterate	primary	secondary	all
rNRSP_below45	12.443**	2.878	5.832	3.982*	-0.016	0.455	5.285	0.515
	(6.242)	(2.956)	(4.248)	(2.289)	(2.987)	(2.605)	(4.187)	(1.774)
rNRSP_4549	4.674	4.133**	4.145*	3.764***	-0.731	0.643	3.445	0.374
	(4.757)	(1.848)	(2.427)	(1.381)	(2.348)	(2.008)	(3.145)	(1.331)
rNRSP_5054	7.430*	1.559	4.032*	2.876**	-1.092	0.184	-0.286	-0.871
	(4.252)	(1.885)	(2.366)	(1.342)	(2.188)	(2.095)	(3.159)	(1.334)
rNRSP_5559	4.840	1.713	4.889**	2.902**	-1.796	1.494	0.435	-0.428
	(3.846)	(1.748)	(2.486)	(1.331)	(1.987)	(2.092)	(3.521)	(1.315)
rNRSP_6064	2.983	2.056	2.336	1.813	-2.196	1.269	3.976	-0.368
	(3.682)	(1.786)	(2.704)	(1.362)	(1.967)	(2.071)	(4.081)	(1.320)
rNRSP_6569	4.994	1.025	3.542	1.298	-3.224	1.656	7.542	-0.692
	(4.194)	(1.853)	(3.259)	(1.470)	(2.133)	(2.325)	(5.617)	(1.468)
rNRSP_7074	0.904	2.130	4.060	1.376	-2.318	1.161	3.419	-0.433
	(4.080)	(2.279)	(4.355)	(1.732)	(2.326)	(3.067)	(9.559)	(1.736)
rNRSP_7579	-5.111	2.820	3.880	-0.240	-4.567*	2.731	-13.250	-2.272
	(4.376)	(2.856)	(8.001)	(2.127)	(2.772)	(5.090)	(16.982)	(2.264)
rNRSP_above80	9.626	11.558**	3.316	9.527***	0.509	17.172**		3.793
	(6.802)	(4.705)	(10.524)	(3.615)	(4.066)	(8.049)		(3.541)
rage	0.777	1.277***	1.526*	1.112***	0.448	0.413	1.115	0.316
	(1.103)	(0.459)	(0.857)	(0.370)	(0.445)	(0.514)	(0.711)	(0.282)
rage_sq	-0.754	-1.579***	-1.813**	-1.360***	-0.703*	-0.891*	-1.692**	-0.705***
	(0.922)	(0.395)	(0.765)	(0.318)	(0.381)	(0.467)	(0.728)	(0.249)
rmarried	0.983	2.320	-7.610	0.133	-1.835	-2.647	5.158	-1.138
	(4.902)	(3.298)	(4.917)	(2.440)	(2.636)	(3.950)	(6.128)	(2.099)
hchild	-0.662	-0.162	-1.042	-0.450	-0.666	0.249	1.597	-0.117
	(1.561)	(0.667)	(0.950)	(0.515)	(0.716)	(0.902)	(1.524)	(0.535)
hhhres	-0.153	-0.135	-0.174	-0.142	0.356	-0.493	0.382	0.037
	(0.649)	(0.327)	(0.345)	(0.224)	(0.307)	(0.352)	(0.546)	(0.212)
lnhhadurbl	0.926**	0.632***	0.069	0.525***	0.409*	-0.124	-0.130	0.153
	(0.369)	(0.220)	(0.302)	(0.161)	(0.226)	(0.281)	(0.498)	(0.167)
y2013	-3.291	-1.713*	-0.715	-1.376**	-0.003	-1.575	1.483	-0.432
	(2.115)	(0.976)	(1.064)	(0.677)	(1.122)	(1.055)	(1.619)	(0.694)
y2015	-2.486	-1.925**	-0.651	-1.385**	-0.730	-0.562	2.607	-0.169
	(2.115)	(0.953)	(1.163)	(0.680)	(1.138)	(1.070)	(1.764)	(0.709)
urban_nbs	-10.329	4.210**	7.602***	5.539***	0.680	7.371***	3.550	5.166***
	(9.745)	(2.019)	(2.074)	(1.438)	(3.962)	(2.817)	(3.106)	(1.814)
_cons	15.469	26.681*	25.773	30.249***	45.043***	57.749***	58.519***	58.532***
	(35.276)	(13.640)	(24.251)	(11.084)	(14.491)	(14.912)	(17.134)	(8.509)
σ_{hrs}	3.191***	3.163***	3.186***	3.178***	3.203***	3.238***	3.239***	3.229***
	(0.023)	(0.010)	(0.013)	(0.007)	(0.010)	(0.010)	(0.016)	(0.006)
$\rho_{NRSP-hrs}$	-0.139	-0.053	-0.101*	-0.068**	0.040	-0.017	-0.037	0.006
	(0.089)	(0.041)	(0.060)	(0.032)	(0.045)	(0.044)	(0.071)	(0.028)
$\rho_{work-hrs}$	-0.197	0.046	-0.097*	-0.032	-0.199***	-0.117**	-0.021	-0.151***
	(0.188)	(0.043)	(0.053)	(0.033)	(0.073)	(0.047)	(0.126)	(0.037)
Observations	2633	9652	6350	18635	9548	8414	3261	21227
Log likelihood	-10697	-44193	-30466	-85569	-37991	-36904	-14805	-90001

ibid.

1.8.3 Asset Quartiles

While educational levels are highly correlated with household income levels, we directly estimate the heterogeneous responses of individuals in different positions of the wealth distribution. We expect that participants in lower quantiles of asset distribution will respond more to the pension programme. We use the values of household durable assets instead of household incomes because it suffers less from the problem of missing values. We divide the sample into four groups based on the quantiles of the values of household durable assets. The categorization is done separately for each survey wave and separately for men and women. The lower quantiles represent lower values of household durable assets. The estimated probabilities of participating in the NRSP for each subgroup are reported in Table A.3 in the Appendix A.

As shown in Tables 1.15 and 1.16, the positive and significant $\rho_{NRSP-work}$ and $\rho_{NRSP-hours}$ for women in the lowest asset quantile suggests that among them, NRSP participants are more likely to current workers who work for longer hours, probably because they have lower levels of household income and savings, and have to rely on labour incomes for old-age support. The negative and significant $\rho_{work-hrs}$ suggests that women staying in the labour force till older age tend to reduce their hours of working. Coefficients of the NRSP participation variables suggest that female workers coming from the lowest asset quantile are more likely to stop working or reduce their hours of working after receiving the pensions. Women in higher asset quantiles and men in all asset quantiles do not respond to the NRSP, probably due to the low level of benefits the scheme offers.

Table 1.15: Effects of NRSP Participation on Working Probability by Asset Quartiles

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
rNRSP_below45	0.000	-0.430	-0.253	0.302	-0.120	0.533**	0.034	0.139
	(.)	(0.443)	(0.408)	(0.403)	(0.298)	(0.262)	(0.219)	(0.207)
rNRSP_4549	0.191	-0.066	-0.099	0.423**	-0.390**	0.190	0.071	0.130
	(0.219)	(0.237)	(0.236)	(0.206)	(0.173)	(0.153)	(0.148)	(0.139)
rNRSP_5054	0.187	0.048	-0.035	0.289	-0.316*	0.006	-0.175	0.097
	(0.210)	(0.208)	(0.221)	(0.193)	(0.161)	(0.142)	(0.137)	(0.137)
rNRSP_5559	0.222	-0.038	-0.000	0.278	-0.254*	0.048	-0.202	0.100
	(0.191)	(0.196)	(0.201)	(0.193)	(0.143)	(0.134)	(0.134)	(0.140)
rNRSP_6064	0.115	-0.080	0.056	0.255	-0.307**	0.056	-0.011	0.267*
	(0.176)	(0.186)	(0.208)	(0.196)	(0.135)	(0.134)	(0.134)	(0.141)
rNRSP_6569	0.077	-0.103	0.079	0.127	-0.193	0.111	-0.131	0.152
	(0.175)	(0.187)	(0.209)	(0.204)	(0.133)	(0.142)	(0.145)	(0.159)
rNRSP_7074	0.075	-0.051	-0.004	0.477**	-0.304**	-0.035	-0.231	0.143
	(0.178)	(0.198)	(0.221)	(0.231)	(0.137)	(0.152)	(0.159)	(0.185)
rNRSP_7579	-0.130	-0.202	-0.326	0.427	-0.450***	0.107	-0.091	0.096
	(0.189)	(0.219)	(0.260)	(0.260)	(0.152)	(0.188)	(0.195)	(0.246)
rNRSP_above80	0.037	0.059	-0.234	1.484***	-0.105	0.312	-0.388	0.164
	(0.222)	(0.287)	(0.321)	(0.401)	(0.184)	(0.267)	(0.274)	(0.367)
rage	0.067**	0.099**	0.029	0.112**	0.077***	0.060**	0.042	0.015
	(0.034)	(0.046)	(0.046)	(0.057)	(0.030)	(0.029)	(0.032)	(0.037)
rage_sq	-0.091***	-0.128***	-0.069*	-0.143***	-0.101***	-0.091***	-0.072***	-0.059*
	(0.026)	(0.037)	(0.038)	(0.047)	(0.024)	(0.024)	(0.027)	(0.032)
rprimary	0.075	-0.043	0.070	-0.090	-0.003	0.143**	-0.010	0.080
	(0.064)	(0.070)	(0.070)	(0.078)	(0.065)	(0.062)	(0.061)	(0.062)
rsecondary	0.043	0.027	-0.012	-0.135	-0.146	-0.019	-0.035	0.099
	(0.078)	(0.082)	(0.079)	(0.087)	(0.100)	(0.080)	(0.071)	(0.068)
rhighabove	-0.032	0.014	-0.003	-0.123	0.214	0.090	0.127	-0.073
	(0.126)	(0.123)	(0.114)	(0.107)	(0.179)	(0.155)	(0.138)	(0.104)
rmarried	0.368**	0.462*	0.106	0.497	0.132	0.263	0.299	0.534**
	(0.173)	(0.256)	(0.259)	(0.326)	(0.130)	(0.163)	(0.204)	(0.226)
hchild	0.038	0.112**	-0.024	0.181***	0.029	0.049	-0.014	0.036
	(0.045)	(0.050)	(0.057)	(0.064)	(0.039)	(0.043)	(0.051)	(0.054)
hhhres	0.030	-0.053*	0.058**	0.025	0.002	0.003	-0.027	0.018
	(0.023)	(0.027)	(0.029)	(0.026)	(0.020)	(0.022)	(0.021)	(0.022)
lnhhadurbl	-0.016	-0.043	0.107	-0.046	0.011	-0.140**	-0.093	-0.071*
	(0.013)	(0.078)	(0.109)	(0.051)	(0.011)	(0.063)	(0.087)	(0.040)
y2013	-0.096	-0.015	0.093	-0.206**	0.136**	-0.002	0.140*	-0.065
	(0.089)	(0.100)	(0.096)	(0.091)	(0.067)	(0.073)	(0.071)	(0.068)
y2015	-0.092	-0.008	0.017	-0.113	0.040	0.030	0.015	-0.019
	(0.087)	(0.102)	(0.104)	(0.095)	(0.069)	(0.075)	(0.075)	(0.072)
urban_nbs	-0.403**	-0.962***	-0.949***	-0.637***	-0.900***	-0.874***	-0.762***	-0.920***
	(0.162)	(0.163)	(0.140)	(0.132)	(0.145)	(0.133)	(0.124)	(0.105)
_cons	-0.730	-0.360	-0.096	-0.092	-0.429	0.600	1.160	2.108*
	(1.105)	(1.539)	(1.641)	(1.827)	(0.958)	(0.970)	(1.188)	(1.190)
$\rho_{NRSP-work}$	0.042	0.130	0.077	-0.147	0.228***	-0.004	0.129*	0.014
	(0.101)	(0.110)	(0.121)	(0.110)	(0.080)	(0.076)	(0.076)	(0.077)
$\rho_{work-hrs}$	-0.120*	-0.097	0.082	-0.027	-0.135*	-0.149	-0.175***	-0.156***
	(0.072)	(0.135)	(0.058)	(0.063)	(0.082)	(0.133)	(0.062)	(0.054)
Observations	4664	4732	4666	4573	5312	5383	5262	5269
Log likelihood	-19968	-21893	-21797	-21593	-20454	-23038	-23248	-22918

Table 1.16: Estimated Effects of NRSP Participation on Working Hours by Asset Quartiles

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
rNRSP_below45	-3.708	9.996**	2.734	4.989	-4.576	2.128	-1.593	2.915
	(6.362)	(4.262)	(3.963)	(4.347)	(4.444)	(3.483)	(3.218)	(3.352)
rNRSP_4549	7.815**	6.602**	1.423	1.943	0.734	1.752	-2.013	1.566
	(3.900)	(2.711)	(2.386)	(2.669)	(3.295)	(2.674)	(2.557)	(2.643)
rNRSP_5054	6.216	3.511	1.837	1.260	-2.862	2.473	-2.289	-0.896
	(3.847)	(2.611)	(2.370)	(2.635)	(3.243)	(2.690)	(2.623)	(2.665)
rNRSP_5559	5.496	2.268	1.798	2.103	-4.700	3.053	-1.650	1.075
	(3.714)	(2.551)	(2.344)	(2.708)	(2.947)	(2.562)	(2.544)	(2.680)
rNRSP_6064	4.359	1.198	1.577	-0.518	-2.560	0.541	-0.352	0.304
	(3.594)	(2.579)	(2.421)	(2.936)	(2.730)	(2.588)	(2.600)	(2.868)
rNRSP_6569	3.461	-1.395	0.912	3.742	-5.536*	0.799	3.534	-1.376
	(3.688)	(2.759)	(2.737)	(3.355)	(2.845)	(2.983)	(2.860)	(3.407)
rNRSP_7074	3.725	-1.315	0.823	1.548	-4.399	0.209	1.881	3.630
	(3.857)	(3.264)	(3.453)	(4.838)	(3.246)	(3.330)	(3.579)	(4.753)
rNRSP_7579	0.120	-5.096	-0.822	6.025	-1.635	-6.072	-2.383	4.111
	(4.115)	(4.331)	(4.967)	(6.147)	(4.201)	(4.155)	(4.188)	(6.971)
rNRSP_above80	9.985	3.024	10.165	15.975*	1.432	-3.017	20.218***	-6.359
	(6.137)	(6.920)	(10.735)	(8.431)	(5.752)	(6.003)	(7.361)	(11.604)
rage	0.646	1.037	1.598**	1.569*	0.846	0.081	0.632	0.083
	(0.550)	(0.807)	(0.815)	(0.944)	(0.653)	(0.505)	(0.503)	(0.555)
rage_sq	-0.922**	-1.110	-1.876***	-1.822**	-1.067*	-0.457	-1.048**	-0.557
	(0.456)	(0.703)	(0.720)	(0.832)	(0.558)	(0.465)	(0.446)	(0.500)
rprimary	0.216	-0.748	-1.438	-0.602	1.270	-1.847	-0.261	0.258
	(1.084)	(1.021)	(1.052)	(1.212)	(1.409)	(1.131)	(1.169)	(1.198)
rsecondary	0.336	-0.360	-1.793	-2.108*	1.808	0.813	1.801	-1.106
	(1.335)	(1.175)	(1.140)	(1.258)	(1.973)	(1.484)	(1.399)	(1.345)
rhighabove	1.962	-0.234	-4.560***	-1.992	2.679	2.705	1.633	0.943
	(2.270)	(1.751)	(1.621)	(1.557)	(3.318)	(2.588)	(2.517)	(2.159)
rmarried	2.003	2.217	4.003	-15.126***	1.966	-3.520	-2.393	-3.168
	(4.449)	(5.356)	(4.636)	(5.676)	(3.202)	(4.835)	(4.401)	(5.096)
hchild	-0.690	0.192	-0.410	-1.470	-0.516	-0.618	0.142	0.770
	(1.079)	(1.053)	(0.993)	(1.118)	(0.984)	(0.988)	(1.113)	(1.217)
hhhres	-0.090	-0.205	-0.225	-0.097	0.466	0.290	-0.069	-0.302
	(0.499)	(0.469)	(0.486)	(0.425)	(0.473)	(0.488)	(0.450)	(0.450)
lnhhadurbl	0.670**	1.554	0.074	0.808	0.245	0.654	-1.728	2.166**
	(0.281)	(1.272)	(1.602)	(0.751)	(0.288)	(1.310)	(1.660)	(0.898)
y2013	-1.795	0.291	-1.447	-2.566**	-0.389	1.217	-0.058	-2.636**
	(1.868)	(1.507)	(1.300)	(1.228)	(1.551)	(1.504)	(1.445)	(1.318)
y2015	-0.256	-1.705	-0.704	-2.941**	1.521	1.013	-0.770	-2.566*
	(1.846)	(1.506)	(1.377)	(1.257)	(1.579)	(1.528)	(1.510)	(1.362)
urban_nbs	2.060	7.136**	4.690**	7.599***	4.966	7.684*	3.204	7.578***
	(4.500)	(3.523)	(2.138)	(2.087)	(4.823)	(4.091)	(3.212)	(2.473)
_cons	39.786**	21.025	21.048	16.974	37.695*	60.952***	61.105***	55.826***
	(17.530)	(24.848)	(26.789)	(28.187)	(19.956)	(16.248)	(20.116)	(18.920)
$ln\sigma_{hrs}$	3.196***	3.190***	3.151***	3.153***	3.222***	3.211***	3.229***	3.228***
	(0.016)	(0.014)	(0.014)	(0.014)	(0.013)	(0.014)	(0.011)	(0.012)
$\rho_{NRSP-hrs}$	0.042	0.130	0.077	-0.147	0.228***	-0.004	0.129*	0.014
	(0.101)	(0.110)	(0.121)	(0.110)	(0.080)	(0.076)	(0.076)	(0.077)
$\rho_{work-hrs}$	-0.120*	-0.097	0.082	-0.027	-0.135*	-0.149	-0.175***	-0.156***
	(0.072)	(0.135)	(0.058)	(0.063)	(0.082)	(0.133)	(0.062)	(0.054)
Observations	4664	4732	4666	4573	5312	5383	5262	5269
Log likelihood	-19968	-21893	-21797	-21593	-20454	-23038	-23248	-22918

1.9 Robustness Check

1.9.1 Estimating Using Rural Communities

In this section, we re-estimate the trivariate models using the subgroup of rural residents locating in rural communities, and check whether the conclusions are consistent with those from the main discussions. Urban communities have smaller groups of rural residents eligible to participate in the NRSP. Rural residents living in urban communities may differ systemically in terms of labour supply behaviour from their counterparts staying in rural communities. They are more likely to be migrant workers not having local residence permits, or urban hukous, and are more likely to take physically-demanding, non-agricultural jobs. According to the pension policy, rural residents should enrol in the NRSP and contribute to their pension accounts in their registered locations. Therefore, migrant workers, especially the age-ineligible among them, are less willing to participate than local workers do. Li, Wang and Zhao (2018) do not differentiate between rural and urban communities, but they do drop individuals living in places other than their registered location. The identification of rural communities is based on the official categorization of National Bureau of Statistics of China.

Table 1.17 shows the trivariate estimation outputs using only the sample of rural residents not formally retired and living in rural areas, separately for men and women. Compared to Table 1.7, the excluded variable of the community-level placement of the NRSP shows a higher level of significance and a larger magnitude, meaning that rural residents living in rural areas are more likely to enroll in the NRSP compared to those living in urban areas. Correlations between unobserved effects that predict different outcomes show similar signs and magnitudes as they are in models using the full sample.

Compared to the estimations for the whole sample in Table 1.7, the signs and magnitudes of the NRSP variables do not change much, but the significant levels drop, probably due to the shrinking sample size.

Table 1.17: Trivariate Estimation Using Sample from the Rural Communities

	Males			Females		
	(1) NRSP	(2) Work	(3) Hours	(4) NRSP	(5) Work	(6) Hours
NRSP in place	1.997***			2.012***		
	(0.061)			(0.058)		
rNRSP_below45		0.049	2.738		0.138	1.909
		(0.251)	(2.564)		(0.134)	(2.042)
rNRSP_4549		0.178	3.707**		-0.007	0.871
		(0.125)	(1.583)		(0.088)	(1.556)
rNRSP_5054		0.167	2.387		-0.090	-0.529
		(0.114)	(1.539)		(0.087)	(1.568)
rNRSP_5559		0.166	1.404		-0.108	-0.557
		(0.111)	(1.501)		(0.081)	(1.528)
rNRSP_6064		0.137	0.870		0.004	0.557
		(0.108)	(1.544)		(0.081)	(1.535)
rNRSP_6569		0.074	1.146		-0.023	0.048
		(0.106)	(1.662)		(0.082)	(1.689)
rNRSP_7074		0.085	-0.290		-0.134	0.971
		(0.112)	(1.873)		(0.089)	(1.963)
rNRSP_7579		-0.077	-1.208		-0.107	-0.796
		(0.123)	(2.281)		(0.107)	(2.518)
rNRSP_above80		0.196	8.962**		0.098	5.678
		(0.151)	(3.819)		(0.143)	(3.859)
rage	0.100***	0.073***	1.100***	0.090***	0.073***	0.694**
	(0.014)	(0.026)	(0.425)	(0.011)	(0.018)	(0.321)
rage_sq	-0.075***	-0.099***	-1.319***	-0.067***	-0.099***	-1.027***
	(0.011)	(0.021)	(0.364)	(0.009)	(0.015)	(0.286)
rprimary	0.053	0.025	-0.864	0.017	0.070*	-0.963
	(0.033)	(0.045)	(0.634)	(0.035)	(0.042)	(0.743)
rsecondary	0.040	-0.059	-1.195*	0.055	0.000	0.213
	(0.038)	(0.052)	(0.719)	(0.045)	(0.054)	(0.913)
rhighabove	0.087	0.041	-0.717	0.078	0.018	3.727**
	(0.055)	(0.075)	(1.083)	(0.076)	(0.096)	(1.793)
rmarried	0.255*	0.309**	-0.116	-0.123	0.209**	-0.832
	(0.135)	(0.121)	(2.741)	(0.111)	(0.091)	(2.260)
hchild	0.078***	0.112***	-0.138	0.075***	0.018	-0.264
	(0.027)	(0.026)	(0.565)	(0.027)	(0.023)	(0.597)
hhhres	-0.022*	0.007	0.023	-0.001	-0.002	0.036
	(0.013)	(0.013)	(0.250)	(0.012)	(0.011)	(0.236)
lnhhadurbl	0.028***	0.000	0.519***	0.015*	0.004	0.034
	(0.010)	(0.009)	(0.175)	(0.009)	(0.008)	(0.184)
y2013	0.567***	-0.107**	-0.671	0.542***	0.028	-0.642
	(0.034)	(0.054)	(0.818)	(0.032)	(0.041)	(0.838)
y2015	0.509***	-0.145***	-0.215	0.455***	-0.026	-0.291
	(0.035)	(0.056)	(0.817)	(0.033)	(0.042)	(0.853)
_cons	-5.284***	-0.741	29.132**	-4.749***	-0.544	43.759***
	(0.448)	(0.820)	(12.878)	(0.372)	(0.564)	(9.655)
$ln\sigma_{hrs}$	3.169***	3.169***	3.169***	3.220***	3.220***	3.220***
	(0.008)	(0.008)	(0.008)	(0.007)	(0.007)	(0.007)
$\rho_{NRSP-work}$	0.012	0.012	0.012	0.094**	0.094**	0.094**
	(0.059)	(0.059)	(0.059)	(0.046)	(0.046)	(0.046)
$\rho_{NRSP-hrs}$	-0.071*	-0.071*	-0.071*	-0.001	-0.001	-0.001
	(0.036)	(0.036)	(0.036)	(0.034)	(0.034)	(0.034)
$\rho_{work-hrs}$	-0.086	-0.086	-0.086	-0.109*	-0.109*	-0.109*
	(0.066)	(0.066)	(0.066)	(0.057)	(0.057)	(0.057)
Observations	14063	14063	14063	15696	15696	15696
Log likelihood	-64978	-64978	-64978	-67945	-67945	-67945

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, living in the rural areas.

1.9.2 Panel Data Estimation

We re-estimate the trivariate models in the previous sections using alternative fixed-effects (FE) approaches to relieve the concern of time-varying confounders. We estimate fixed-effects instrumental variable (FE IV) models separately for the working decision and the hours of work decision. The community-level placement of the NRSP is used as an instrumental variable for the potentially endogeneous variable of participating in the NRSP. To account for the sample selection bias when we study the working hours of the current workers, we use the panel heckit approach. Specifically, we estimate a random effects Probit model for the probability of working, in which being currently working is a function of all covariates. We construct the inverse mill ratio and include it as an extra independent variable in the FE IV estimation of hours of work. A significant effect of the inverse mill ratio has the same indication as a significant $\rho_{work-hrs}$ in the trivariate model does, and shows that there are unobserved effects that predict both working probability and hours of working. Not accounting for the unobserved effects and estimating a linear model of working hours cause biased estimation of the covariates.

Although the FE model controls for unobserved, time-varying individual characteristics, it has several drawbacks. First of all, it does not account for the binary nature of the endogenous, pension participation status. The two-stage, FE IV model also does not account for the binary nature of the working status, and estimates a linear model. Secondly, multi-stage estimation is less efficient than maximum likelihood estimation, though it is less computationally demanding, and can impute uncorrective standard errors. We report the FE estimation outputs in the Appendix A.

1.9.3 Estimating Intention-To-Treat Effects

We estimate the intention-to-treat (ITT) effect of NRSP-eligibility, or the average treatment effects across individuals actually enrolled and not enrolled in the NRPS in communities that have started the NRSP, by replacing the possibly endogenous, individual participation status with the exogenous community-level placement of the NRSP. Instead of estimating a FE IV model, we can estimate a single FE model of working probability or working hours of the eligible NRSP participants. Apart from the single variable of community-level placement of the NRSP, we use an alternative variable of community-level duration of the NRSP programme to estimate a second model of ITT effects of NRSP-eligibility. The assumption is that the longer the NRSP has been in place in the local communities, the more knowledge and better commitment the eligible rural residents will build upon the scheme, and the more likely that they will participate. The community-level duration of the NRSP programme is constructed as the difference between the year a community started the NRSP and a given survey year. We compare the signs and significant levels of ATE and ITT estimation outputs in the Appendix. FE IV estimates tend to have similar signs and significant levels as the FE ITT estimates. Nonetheless, FE ITT estimates overall have smaller magnitudes compared with FE IV estimates.

1.9.4 Alternative Definition of Working Status

We use an alternative definition of working status by categorizing only individuals who work for more than 2 hours per week as workers and individuals who work for 2 hours or less as non-workers or the economically inactive. Similarly, for the hours of work, we only keep the reported weekly hours if they are above 2 and set the rest as missing. We take the log of weekly working hours to study the percentage change. FE estimations in the Appendix A all use the newly defined working status and hours of work as dependent variables. The conclusion remain unchanged compared to our main disccusions.

1.10 Conclusion

The paper studies the labour supply responses of rural workers after they participate in the New Rural Social Pension (NRSP) scheme in China. Focusing on heterogeneity across age groups, our trivariate models with the whole sample predict that male participants overall do not change their working probability, while female participants are less likely to work across all age groups, although the effects are only marginally significant for those aged above 50 and those aged above 70. For participants staying in the labour force, men who are below the pension eligible age of 60 and need to contribute to their pension accounts significantly increase their weekly working hours by 5.86% to 7.67%. Heterogeneity analyses show that the effect mainly comes from age-ineligible, non-agricultural workers. In contrast, women aged above 50 reduce their hours of work, although the effects are not significant.

Heterogeneity analyses are performed on workers from different job sectors, educational groups, and asset quantiles. Male participants aged above 65 are significantly more likely to do non-agricultural jobs instead of agricultural jobs. Female participants aged below 50 are more likely to take non-agricultural jobs. Age-ineligible female participants having completed secondary school or higher levels of education are more likely to be working. Nonetheless, female participants in the lowest quantile of household durable asset values are found to be more likely to stop working after 45. Among the current workers, age-ineligible male participants significantly increase their non-agricultural working hours. The effect also echoes with the estimated increase in working hours for those having completed secondary school or higher levels of education among them. The positive effect on labour intensity is not significant for female workers, regardless of the job sectors, educational groups, and asset quantiles they are in. Estimated correlations between unobserved effects that predict participating in the NRSP and labour supply behaviours suggest that agricultural workers, workers not receiving any formal education, and workers in the lower asset quantiles are more likely to participate in the NRSP than non-workers are, and non-agricultural workers are the least likely to enroll in the NRSP. Among the current workers, male participants, especially those taking non-agricultural jobs and receiving secondary or higher levels of education, coming from the higher asset quantiles, tend to work for less hours than non-participants do.

Studying the contribution decisions of the age-ineligible participants and their labour supply behaviours, we find that contributing above the minimum requirement level does not significantly reduce the working probability of either men or women, but is related to a 15.97% decline in women' weekly working hours, probably due to the expectation of receiving higher pensions after 60.

The negative effects of pension participation on the working probability of female participants, especially those from the low-income groups, and the negative effect on weekly working hours of females coming from low-educated, low-income groups and working in the agricultural sector, though not very significant, suggest that an increase in pension benefits and expected SSW substitutes for current labour income and results in female workers reducing their hours of work or completely stopping to work. The negative labour supply responses also differ across age groups, and older contributors tend to respond more than younger contributors do. This can be explained by a life-cycle pattern of the pension effect, as our theoretical models predict that older workers approaching the pension-eligible age have less uncertainty about the future and the survival risk, and need to contribute for less years than the younger participants do. The lack of substitution effect on male pensioners' working probability or working hours can be explained by the low levels of basic benefits that the NRSP offers. It might only see behavioural changes in the low-income groups and workers in poor health status or workers suffering from high opportunity cost of working, and those making high contributions to the pension scheme. The gender difference in labour supply behaviour can also be explained by the fact that labour suppy is a family decision and men tend to be the main income earners and are thus less likely to exit the labour market.

Future researches can further investigate the labour supply responses from part-time and full-time workers using multinomial Probit or logit model[10]. Specifically, each agent chooses from the alternatives of not working, part-time work and full-time work, or a a more complex alternative set of not working, part-time agricultural work, part-time non-agricultural work, full-time agricultural work and full-time non-agricultural work. As shown in the Appendix A, the distribution of the logged hours

[10] Rust and Phelan (1997) and Shu (2018) define working status based on weekly working hours: full-time work=weekly hours>30, part-time work=2 $<$weekly hours<30, retirement<2.

of work is not normal, skewing heavily to the right. Part-time workers may behave in a systemically different way from full-time workers, and the difference cannot be identified based on our estimations that pool all workers together and estimate an average treatment effect on working hours. Heteroskedasticity can also compromise the consistency of ML estimation of censored models or truncated models.

Our study on the NRSP provides policy implications on developing countries that are reforming or expanding their non-contributory public transfer programmes to cover the larger population of informal workers. It shows that workers will respond to the pension programs by adjusting their labour supply and timing their retirement. The program also improves the welfare levels of older workers in terms of leisure hours and pension income.

To improve on the current empirical specification, we can account for spouses' characteristics, especially joint labour and joint participation in the pension programme that can significantly affect the individual labour supply decision, participation decision, and contribution decision. Studies estimating structural models of joint exit from the labour market of couples have provided evidence of the key role that complementarity of leisure plays in individual labour participation decisions (Hurd, 1990; Gustman and Steinmeier, 2005). We can also develop a structural model that estimates simultaneous working probabilities of both husband and wife.

We should have also controlled for time-varying community characteristics, especially the local labour market conditions over time. Although only having access to one wave of community survey data, we can make use of administrative data and other survey data. One problem is that we can only identify the sampled communities at the prefecture level.

Chapter 2

Retirement Effect on Cognitive Functioning and Depression Risk of Formal Workers

Population aging has become a growing challenge for both developed and developing countries with the rising life expectancy and declining fertility. The consequential burden of national pension systems and labour force shortages have induced countries to raise normal retirement ages or pension-eligibility ages. Postponing retirement helps stabilize the pension system by delaying or reducing the pension benefits to be paid, and increasing tax or contribution base. The implications and evaluations of these pension reforms relate closely to the effect of retirement on individual and household welfare and behaviours. The literature of retirement or ageing economics has focused on monetary outcomes, or more generally old-life support from various sources including savings, private and public transfer. Non-monetary outcomes, especially individual physical health, health care use, and health behaviour, have also been widely studied. Less attention has been paid to mental health conditions of retirees, including their cognitive decline process, subjective well-being and especially depression symptoms.

There are several reasons why it is important to study mental health of older workers. First of all, serious forms of cognitive impairment can cause disability and dementia in old age (Coe et al., 2012). Severe depression is also associated with disability, medical illnesses such as heart diseases, diabetes (Drew et al., 2010; Fruehwirth, Iyer and Zhang, 2019). Neurological studies find that depressive symptoms in older persons are associated with cognitive declines and a higher risk of developing Alzheimer Disease (AD) (Wilson et al., 2002). Epidemiologic and clinical studies provide evidence on a reciprocal, potentially spiral, relationship between depression and disability in later years (Bruce, 2001), especially for those aged above

70 and experiencing high level of depressive symptoms, with modest differences by gender (Barry et al., 2009).

In addition, mental health problems can affect individual behaviours and cause adverse outcomes such as bad financial decisions and retirement planning that will undermine individual late-life supports and generate significant economic costs. The costs come from public transfers, medical and health care expenditures, the opportunity costs for families and friends taking informal care and the let-go economic benefits of healthy, older population. Older people in good health conditions can make valuable contributions to society by means of consumption, or unpaid labour such as caring for grandchildren and volunteering. It is estimated that healthcare expenditures are about four times higher for the depressed than the non-depressed (Hewlett and Moran, 2014). Therefore, the economic costs have become a growing concern for countries facing an ageing population.

Nonetheless, mental health problems are preventable if we have a better understanding of the risk factors. This paper focuses on the effect of retirement, including both the short-run effect of transition into retirement and the cumulative effect of retirement duration on mental health of retirees.

Identifying the retirement effect on subjective well-being can be challenging as retirement can be a choice and closely related to individual health conditions before retirement. One solution is to find an instrument that is correlated with the retirement decision but not with cognition or depression. Early researches studying European data have used cross-country variations in the eligibility ages for early and normal retirement benefits in Europe as IVs and data from the Survey of Health, Ageing and Retirement in Europe (SHARE) (Rohwedder and Willis, 2010; Coe and Zamarro, 2011; Bonsang, Adam and Perelman, 2012; Mazzonna and Peracchi, 2012). Studies of the US use within-country variation in early and normal pension-eligible ages as IVs and data from Health and Retirement Studies (HRS) (Kofi Charles, 2004; Neuman, 2008; Coe et al., 2012).

Policy reforms, such as pension or Social Security reforms that generate exogenous variations in the retirement probability have also been exploited for identification in some studies (Kofi Charles, 2004; Battistin et al., 2009; Eibich, 2015; Grip, Lindeboom and Montizaan, 2011; Shai, 2018; Atalay, Barrett and Staneva, 2019). The assumption is that these pension policies are unlikely to be set based on the observed age patterns in cognitive functioning and mental health, and thus serve as valid instruments for retirement (Rohwedder and Willis, 2010). These reforms are strong instruments because they affect individual retirement decisions by changing the implicit tax rate on labour earnings (Gruber and Wise, 2005).

This paper examines both the short-run effect of entering formal retirement for formal workers reaching the mandatory retirement age, and the cumulative effect of years in retirement for these retirees on their cognitive functioning, depression risks, and physical well-being. We identify formal retirement using mandatory retirement ages for blue-collar and white-collar workers, and separate analyses for men and women. Data from a nation-wide, longitudinal household survey is used and the sample is restricted to formal workers or formal retirees who have urban residence permits (hukous) and live in urban areas. A bivariate random effects model is used to study changes in mental and physical health outcomes after individual transition into retirement and with years in retirement. For comparison, a simple random effects model is also estimated to compare cognitive functioning and mental health of retirees and current workers, and to compare between retirees having stayed for different years in retirement. To account for the potential state-dependence of the health outcomes, we estimate Arellano and Bond (AB) models with instrumental variables and report the estimation results in Appendix B.

We find a substantial and positive effect of transition into formal retirement on men's scores in all cognitive tests, but a negative effect on women's scores in the mental intactness test and no effect on women's memory tests. Retirement is endogenous to men's scores in cognitive tests as there are unobserved effects that significantly predict lower scores in cognitive tests and the probability of retirement. The positive short-run effects of retirement disappear for men with the number of years in retirement and the negative effect reverses for women. The gender-specific difference in the pattern of cognitive decline after retirement may be explained by the types of social activities they take up and the younger retirement ages for women. There is no significant retirement effect on either men's or women's subjective or physical well-being, as measured by the the number of self-reported depressive symptoms and the number of self-reported physical problems. Retirement is endogenous to physical health of both men and women as there are unobserved effects that significantly predict more physical health problems and the probability of retirement. Risks of depression and physical frailty are not found to significantly change with the years of formal retirement.

The paper contributes to enrich the understanding of retirement effects on individual outcomes and adds to the limited evidence of the retirement effect on mental health based on the context of a developing country. Similar studies have focused on the labour market in the US or European countries, where retirement ages tend to be older and retirement pensions are more generous. Their findings may not apply to the context of developing countries. Moreover, these studies provide mixed empirical evidence because they use different identification strategies and data sets, study different demographic and occupational groups, and investigate different effects (transitional versus progressive). Some find significant and substantial negative effects on cognitive functioning (Rohwedder and Willis, 2010; Mazzonna and Peracchi, 2012; Bonsang, Adam and Perelman, 2012; Celidoni, Dal Bianco and Weber, 2017; Atalay, Barrett and Staneva, 2019) and depression (Grip, Lindeboom and Montizaan, 2011). Others find an overall null effect (Coe and Zamarro, 2011; Coe et al., 2012;

Messe and Wolff, 2019), or a preservative effect of retirement on mental health (Kofi Charles, 2004; Neuman, 2008; Bound and Waidmann, 2007; Insler, 2014; Eibich, 2015; Kolodziej and García-Gómez, 2019).

In addition, studies using cross-country variations in the pension-eligibility ages as instruments in models using cross-sectional data are subject to the potential bias that individuals from different countries can face with different institutional settings and cultures that are correlated with the retirement schemes. Bonsang, Adam and Perelman (2012) point out that Northern-European countries having better health outcomes also set a higher eligibility age for retirement. Bingley and Martinello (2013) argue that cross-country variations in pension-eligibility ages are correlated with differences in years of schooling, which will affect cognitive functioning at old ages[1]. Studies using data from the US would need to consider the change in health insurance upon retirement (Eibich, 2015).

In the case of China, the relatively young retirement age of 50 for the majority of female workers and 60 for most male workers means that with the growing life expectancy, retirees in China now spend more time on retirement and in home environments. The only evidence of a causal effect of retirement on cognitive decline is from Lei and Liu (2018), in which the main analyses are based on cross-sectional data (although they compare with FE IV estimation outcomes in robustness checks), and do not account for the state dependence of health outcomes, the binary nature of the endogenous variable of retirement status. They also do not look into outcomes of subjective well-being and physical health, or the progressive effect of retirement.

Methodologically, we contribute to the literature by adopting the bivariate random effects model. It accounts for the binary nature of the endogenous retirement status by using probit random effects model to estimate the retirement transition. It also accounts for the individual unobserved heterogeneity that determines both health outcomes and retirement by allowing for the random effects of the two equations to correlate with each other and share a zero-mean, bivariate joint normal distribution. The Arellano and Bond model with an endogenous variable is estimated to account for potential state-dependence of the health outcomes. Only few of the previous studies consider the state dependence of health outcomes, and none of them use the dynamic model to address the concern that state-dependence cannot be fully accounted for in mixed-effects models. Insler (2014) use changes in health status instead of health levels as the dependent variable to avoid the simultaneity problem. Dave, Rashad and Spasojevic (2006) stratify the sample to include only individuals who have no major illness in waves before retirement, and no reported worsening of health between waves before retirement. Nonetheless, state-dependence is generally not significant for scores in cognitive tests or depressive symptoms (Appendix B), and the main discussions are

[1] Using simulation and a replication of Rohwedder and Willis (2010)'s study, they show that models not controlling for schooling produce negatively biased estimates of retirement effect.

based on bivariate random effects models that do not account for state-dependence.

The paper is organized as follows. Section 2.2 summarizes the related literature of retirement and cognitive functioning or mental health. Section 2.3 introduces the institutional background of mandatory retirement policy for formal workers in China. Section 2.4 explains some identification issues and the empirical strategies we use to identify the effect of formal retirement. Section 2.5 introduces the data and the sample, the outcomes variables and the explanatory variables we use for the estimation, and presents the summary statistics. Section 2.6 specifies the econometric model, followed by the empirical results in Section 2.7. Section 2.8 provides some robustness checks, and section 2.9 concludes and discusses the policy implications of our findings.

2.1 Relevant Literature

2.1.1 The Underlying Mechanisms between Cognitive Ageing, Cognitive Decline and Retirement

This section introduces factors that can affect the cognitive decline process and focuses on retirement as an important factor that affects cognitive decline. The mechanisms underlying the retirement effect on cognition include changes in lifestyle and health behaviours (Coe and Zamarro, 2011; Mazzonna and Peracchi, 2012; Rohwedder and Willis, 2010), the loss of intellectual stimulation in job (Salthouse, 1996; Hultsch et al., 1999), reduced investment in cognitive repair activities (Grossman, 1972; Stine-Morrow, 2007; Rohwedder and Willis, 2010; Mazzonna and Peracchi, 2012), the loss of social interactions and networking at work environment (Glass et al., 1999; d'Hombres et al., 2010; Börsch-Supan and Schuth, 2013).

Cognitive decline forms a fundamental aspect of the ageing process. Although cognitive-ageing is a natural, inevitable process, the progression of cognitive decline is not uniform or exogenous, but related to risk factors including genetic factors, medical comorbidities, lifestyle, psycho-social factors, and other factors yet to be identified (Coe et al., 2012). If the progression is slow, it may not seriously affect individual well-being.

The economic literature has provided evidence of some factors that can affect the cognitive ageing process. They include education (Banks and Mazzonna, 2012) and family background and parental investment that are highly associated with educational outcomes (Cunha and Heckman, 2007) [2]. Early-life disease may affect brain development and cause impairment throughout life, or speed cognitive decline

[2]Possible mechanisms include that highly-educated individuals are better at maintaining their cognitive ability, that higher early-life cognitive ability persists into old age, and that education reflects other early-life economic advantages not captured by childhood socio-economic status that

in old age (Case and Paxson, 2009). The conceptual framework provided by Stern (2002) suggests that higher levels of cognitive reserve will help individuals slow down the cognitive ageing process.

Studies of retirement and cognitive decline find that retirement can substantially change lifestyle, in aspects of intellectual stimulation, social networks, consumption, physical activity, health behaviours etc. (Coe and Zamarro, 2011; Mazzonna and Peracchi, 2012; Rohwedder and Willis, 2010). Active and socially integrated lifestyle in late life in the social, mental, and physical dimensions has been found to benefit cognition and protect older people against dementia. Zantinge et al. (2013) provide a narrative review of the studies published between 2001 and 2013 on transition to retirement and changes in health behaviours including changes in smoking, alcohol consumption, physical activities, and dietary habits. Only involuntary retirement leads to an increase in alcohol and cigarettes consumption due to adjustment problems and stress. Female retirees spend more time on leisure-time physical activities and have healthier dietary habits compared to employed women. Depression and physical health problems can also affect cognitive decline, as recent psychiatry literature shows that depressive symptoms can be an early manifestation of dementia. High blood pressure and its associated vascular damage and overt brain damage are well-known risk factors for cognitive impairment (Skoog and Gustafson, 2006).

Several hypotheses have been introduced to explain why retirement can cause cognitive decline. The mental exercise hypothesis suggests that continued intellectual stimulation is important in retaining a high level of cognitive functioning (Salthouse, 1996; Hultsch et al., 1999). Retirement from an intellectually demanding job that provides such kind of stimulation without compensatory informal training or studying or cognitively challenging social interactions, can potentially lead to a cognitive decline in old age.

In economics, Grossman (1972)'s health model also predicts that retirees or employees approaching retirement may lack incentives to invest in cognitive repair activities, and reduce investment under the expectation of a declining return to investment in work-related human capital. This can increase the rate of cognitive decline after retirement. Similar conclusion is found in the 'Dumbledore hypothesis' of cognitive ageing in psychology (Stine-Morrow, 2007), and is defined as 'on-the-job retirement hypothesis' in Rohwedder and Willis (2010). Mazzonna and Peracchi (2012) present a discrete-time version of the Grossman (1972) model as a theoretical framework and argue that besides the market incentives that disappear after retirement, non-market incentives, relative prices and discount rates etc. can also affect the amount of repair investment. Empirically, the increasing rate of decline after retirement is captured by years in retirement (Bonsang, Adam and Perelman, 2012; Mazzonna and Peracchi, 2012).

---

the literature controls for (Case and Paxson, 2008).

The social capital literature suggests that social interactions and networking at work environment can buffer individuals from health shocks, and losing them after retirement may have a negative effect on health (Glass et al., 1999; d'Hombres et al., 2010). Börsch-Supan and Schuth (2013) find that the shrinkage of social networks after retirement causes cognitive levels to decline by one-third.

Nonetheless, retirement can improve cognitive functioning and mental health for retirees who are relieved from work-related strains, adopt a healthier lifestyle, and participate in leisure activities that are equally or more intellectually stimulating than their jobs. This is more likely for blue-collar workers who take physically demanding jobs, as the literature has provided evidence for the negative health effects of working in physically demanding occupations (See Ravesteijn, van Kippersluis and van Doorslaer (2013) for a review).

2.1.2 Mental Health and Mechanisms Behind Retirement Effect on Mental Health

Among the empirical studies of retirement effects on mental health or subjective well-being, most of them refer to theoretical frameworks rather than empirical evidence to explain the mechanisms through which retirement affects individual mental health (Eibich, 2015). They argue that retirement is associated with a decline in income, social capital or networks and physical activities in the work environment, a relief from physical or mental strain on jobs (Eibich, 2015). Human capital model by Grossman (1972) predicts that retirees may invest more on health because they have more leisure time and face with lower opportunity cost, or they may invest less if they expect lower returns to health or cognitive preservation investment (retirement pension does not depend on health status). In the psychology literature, the theories predict that retirees who hold bridge jobs, have better retirement planning, have more social interactions other than social networks at work such as interactions with retired spouses and friends, are likely to experience the minimum changes in psychological well-being during the retirement transition. Individuals who retire from physically demanding or highly stressful jobs, or who have low job satisfaction at employment, are more likely to benefit with improved mental health after retirement. Nevertheless, subjective well-being is highly associated with physical health conditions and financial status, and individuals who experience significant health declines or income loss in the retirement transition can suffer from worse mental health. Off-time retirement, either earlier or later than social norms or individual expectations, and unhappy marriage prior to retirement can also cause a higher level of stress (Wang, 2007).

The limited empirical evidence on the underlying mechanisms focus on individual health behaviours, individual time use, and heterogeneity effects. Eibich (2015) finds that retirement improves self-reported health and mental health by 0.25 standard deviation (SD), decreases outpatient care utilization by 0.2 SD. The underlying mechanisms include that retirement reduces work-related mental stress and physical strain, reduces the probability of smoking by about 5%, increases sleep duration by 40min per day, and increases regular physical exercises (e.g. repairs and gardening) by 10%. Insler (2014) also reports an increase in vigorous physical activity and the probability of quitting smoking. Studies looking at heterogeneous effects across occupational groups tend to find that retirement is beneficial to blue-collar workers taking physically-demanding jobs (Coe et al., 2012; Mazzonna and Peracchi, 2017), primarily due to relief from work-related strains. Atalay, Barrett and Staneva (2019) investigates into three domains that retirement can affect cognition: mental exercise, socialisation, and time allocation, and they find that increasing time spent in mental and household activities can explain the relatively modest cognitive decline for women. Nonetheless, the above evidence is only suggestive and relies on the assumption that changes in health behaviour and time use affect physical and mental health. There may be reverse causality between health, health behaviour and time use (Eibich, 2015). Mediation analysis provides evidence that retirement effect on mental and physical health is induced by changes in health behaviour and time use given that the effect of retirement decreases after controlling for the mechanisms (Lei and Liu, 2018; Eibich, 2015).

Heterogeneity of retirement effects can also provide some evidence for the mechanisms behind retirement effect on mental and physical health. Blue-collar workers taking physically-demanding jobs tend to be the only occupational group that reports better cognition or mental health after retirement (Coe et al., 2012; Eibich, 2015). Eibich (2015) finds that mental health only improves for those retiring from straining occupations. Coe and Zamarro (2011) look at the differential retirement effects depending on the age when the individual retired (ages 57–59, ages 60–64, and ages 65–69), and find that people retiring between the ages of 60 and 64 are 54 percent less likely to be in bad health than those retiring before age 57. The age gradient of the retirement effect is also used by Lei and Liu (2018) to explain the gender difference in cognitive decline as women in China retire 5-10 years earlier than men do. However, the gender-specific cognitive decline they find can be due to selection effect as men are more likely to work in government or SOEs than women do (Zhao and Zhao, 2018).

Again, there is inconclusive empirical evidence about the retirement effect on subjective well-being because retirement can be a choice and closely related to individual health characteristics before retirement. Economics studies have provided inconclusive evidence using different identification strategies and measures of health. Studies controlling for an extensive range of confounding factors tend to find a negative health effect of retirement (Dave, Rashad and Spasojevic, 2006), while studies using an IV strategy by exploiting an exogenous variation in social security policy generally find no such negative effect, or even positive health effects (Kofi Charles, 2004; Bound and Waidmann, 2007; Neuman, 2008; Coe and Zamarro, 2011).

2.1.3 Retirement and Health in China

The literature of retirement effect on health summarized above are based on the context of developed countries, and there have been few studies looking at Chinese data. Exploiting the changes in the mandatory retirement ages, Che and Li (2018) find that the probability of reporting fair or poor health status decreases by 34% after retirement among white-collar male workers. They provide evidence that increasing exercise, reducing smoking and drinking behaviours are possible channels behind the retirement effect, and justify the changes in health behaviours with declining cost of adopting health-related activities after retirement. Adopting the similar identification strategy, Chen, Geldsetzer and Bärnighausen (2020) find that retirement increases the stress level by 45% for women but reduces the level by 21% for men.

Using a fuzzy regression discontinuity design at statutory retirement ages, Zhang, Salm and van Soest (2018) find that retirement increases healthcare utilization, the number of functional limitations and chronic diseases (e.g. hypertension, diabetes and stomach disease) and body mass index (BMI). They explain that the reduced opportunity cost of time after retirement drives the increase in inpatient care use and more health problems being diagnosed, and the evidence does not reflect causal effects of retirement on genuine health. Finally, they find that the likelihood of foregoing inpatient care increases by 20% for the low educated group, suggesting that the less-educated retirees tackle the health shocks worse due to an income drop after retirement. The high copayments in healthcare in China still discourage low-income groups from using health care.

Using the same data and methods in Zhang, Salm and van Soest (2018), Li (2017) also finds a 58.6% decline in the probability of reporting satisfactory health, and the depression index increases by 30.2%. Drinking behaviour also declines by 53% for male retirees, probably due to less work-related drinking that is common for male workers in China. No significant retirement effect is found for women. Similarly, Feng and Zhang (2018) find that the provision of grandchild care increases by 29% for women (which translates into 1.7 more hours per day) and 21% for men after the transition to retirement. Their later study finds that retirement increases weight by 2.1 kg and BMI index by 0.92 units for men, mainly driven by low educated men (secondary school or below), but has no impact on women (Feng, Li and Smith, 2020).

Lei et al. (2015) find that social networks have ambiguous effects on subjective well-being (SWB), as it can on the one hand resolve stress, provide social support and information, and on the other hand can also consume resources and be subject to relational constraints. The number of friends is more important than the number of relatives. Marriage, social activities, and participation in groups also matter. Meng and Xue (2020) find that one standard deviation increase in social networks reduces the measured mental health problems of rural migrant workers by 0.47 to 0.66 standard deviation.

2.2 Institutional Background

In urban China, workers in the formal sector are subject to mandatory retirement ages and eligible to receive pensions afterwards. According to the official document issued by the State Council in 1978, normal retirement ages are 60 for men, 55 for female civil servants or managers, and 50 for other female workers. Early retirement at an age five or less years younger than the normal retirement age is allowed for people taking physically-demanding jobs, civil servants working for 30 years or more, workers disabled in jobs after certain years of contribution, and for redundant workers during the 1990s state sector restructuring[3].

Despite the efforts of local governments to encourage reemployment of laid-off workers during the reform of state-owned enterprises (SOEs), the percentage of re-employment declined over time to just 30 percent in 2003. Compensation, subsidy and unemployment insurance were offered to redundant workers [4]. Laid-off workers within five years to the normal retirement age can opt for 'internal retirement' where they received a one-off compensation determined by their salaries and years of service, and then retirement pensions after reaching the normal retirement age. Giles, Park and Cai (2006) using the China Urban Labour Survey find that the labour force participation rate fell by about 20% for men approaching retirement (aged 55-60), and about 15% for women aged between 40 and 50.

The normal retirement age for civil servants and urban workers in SOEs in China is among the youngest around the world because it is first introduced in 1950s when the average life expectancy was low. From 1997 onward, all employees in urban enterprises were subject to the mandatory retirement age. The enforcement is strict in the public sector but is more flexible in the private sector. Delaying retirement up to five years or more is allowed for some occupation groups including professors and advanced researchers, civil servants at a position no less than the county or department level, and professionals with similar titles.

[3]Based on China Labour Statistical Yearbook in 2006 (available in http://www.stats.gov.cn/tjsj/ndsj/2006/indexeh.htm) from the Ministry of Labour and Social Security (MOLSS), 21 million workers, about 60% of the total SOEs workers, were laid-off from SOEs between 1994 and 2005, or 34 million if including collective enterprises.

[4]According to a survey by Garnaut, Song and Yao (2006) of 11 cities across China, redundant workers received compensation equivalent to three years' salary . However, in poorly-performed enterprises more common in provinces with a high concentration of SOEs, such as Liaoning, Heilongjiang, Sichuan, Chongqing and Hebei, workers were laid-off without any compensation.

Retirement pensions provide incentives for retirement for workers in both public and private sectors, and being eligible for receiving pensions can be an important factor in deciding when to stop working, especially for workers in the public sector who receive more generous benefits than others. The first formal pension system was established in 1951 and covered all urban workers in the public sector. It offered defined-benefit pensions with a replacement rate as high as 75-90% of wages. Workers were also entitled to welfare packages including allocated houses, medical insurance and social security.

The current social pension system in China was established during the pension reform in 1997 when the government aimed to lower the replacement ratio and extend the coverage to workers in other sectors. Specifically, Basic Old Age Insurance (BOAI) covers workers formally employed in for-profit enterprises in public or private sectors. Public Employee Pension (PEP) covers civil servants and employees in non-profit government institutions such as schools and public hospitals. The BOAI merged with PEP in 2015 and became a uniform programme for all formal employees in urban sectors. The pension schemes vary in their contribution and benefit rules, with significant inequality between formal and informal workers, and across regions as pensions are managed by provincial governments.

For the BOAI, the target replacement ratio set by the Ministry of Human Resources and Social Security (MOHRSS) is 59.2% of the local average wage. The eligibility ages for retirement benefits are 50 for female blue-collar workers, 55 for female white-collar workers, and 60 for males. Before merging with BOAI in 2015, the PEP was more generous than the other schemes and did not ask for any contribution from public employees. The average replacement ratio was 80-90% of pre-retirement wages. After the PEP was unified with the BOAI in 2015, public employees are subject to the same contribution and benefit rules as formal employees in other sectors. People retired before the 2015 reform are unaffected, while those retiring after 2015 are subject to a transitional arrangement for financing individual accounts. Feng and Zhang (2018) provide a detailed introduction of the pension system in China.

Besides the compulsory BOAI, formal employees may also participate in Enterprise Annuity (EA), an employer-sponsored pension system introduced in 1991. The coverage of EA remains small, and by 2017 it had 23.3 million participants, accounting for about 5.8% of the coverage of BOAI. Only large SOEs tend to offer the Enterprise Annuity scheme. Apart from the retirement pension, urban formal workers are also covered by the more generous Urban Employee Health Insurance Programme and stop paying for insurance premiums after retirement.

2.3 Data and Summary Statistics

2.3.1 The China Health and Retirement Study (CHARLS)

China Health and Retirement Study (CHARLS) is a biannual, nation-wide longitudinal survey initiated to study health and retirement behaviours of the elderly labour aged 45 years or above and their spouses in China. It is equivalent to the Health and Retirement Study (HRS) in the United States. The 2011 baseline survey collects information of 17,708 respondents in 10,257 households residing in 150 counties in 28 provinces[5]. These respondents are followed-up in 2013, 2015, and 2018, though some dropped out in the middle of the surveys and new participants are included to replace them.

We choose the CHARLS because it contains the most detailed and comprehensive information on older people's physical and mental health status. It also provides tests on cognitive functioning and depressive symptoms. The population of individuals aged 45 and above, or the middle-aged and the elder, are ideal to study the relationship between health and retirement, as they are faced with more serious forms of cognitive declines and more severe mental or physical health problems compared with the younger population.

The *Gateway to Global Aging Data* has published the harmonized CHARLS dataset and codebook to facilitate cross-country comparisons. The paper uses data of the Version C of the Harmonized CHARLS dataset and supplements it with some variables from the original dataset that are not appearing in the Harmonized dataset, for example, the detailed information about jobs' characteristics.

2.3.2 Outcome and Control Variables

2.3.2.1 Cognitive Functioning

Cognitive ability tests conducted by CHARLS consist of two words tests and one mental intactness test. In the words test, interviewers read a list of 10 common nouns to the respondents, who were asked both immediately and 5 minutes later, to recall as many of the words as they could in any order. The number of words being immediately recalled and the number of words being recalled 5 minutes later are used to construct measures of memory and the aggregated measure of cognitive functioning. Literature using data from the HRS family surveys to study cognition functioning outcomes have constructed measures of memory in different ways. One stream of literature takes an average of the number of words being immediately recalled and the number of words being recalled 5 minutes later (Smith, McArdle and Willis, 2010; Lei and Liu, 2018),

[5]Tibet is not included in the survey, and two other provinces of Hainan and Ningxia are also not represented in the study due to their small population size.

and the other stream sums up the total number of words being recalled (Rohwedder and Willis, 2010). Some studies look at the two memory test scores separately, but do not account for the correlation between the two (Mazzonna and Peracchi, 2012).

On the one hand, the two memory tests measure different types of memory ability, as the number of words being immediately recalled measures short-term memory and the number of words being recalled 5 minutes later measures episodic memory. Episodic memory can be important for reasoning, and for both fluid and crystallised intelligence. We decide to look at the two outcomes separately. On the other hand, the two types of memory are closely related and the nature of the memory tests means that the number of words being immediately recalled can be used as a state-variable to control for omitted factors that determine both short-term memory and episodic memory, and might be also related to the key retirement variables. Therefore, besides estimating the two memory test scores separately, we also estimate a model that uses the number of words being immediately recalled instead of the lagged dependent variable to control for potential state-dependence of the number of words being recalled five minutes later. Details of the model are specified in the empirical modeling section.

During the test of mental intactness, respondents are asked to answer 10 questions which include successively subtracting 7 from 100 up to five times; naming the date, month, year, day of the week of the survey; redrawing a picture of two overlapped pentagons. A mental intactness index is constructed by summing up the number of correct answers and has a range between 0 and 10. It is used as the third measure of cognitive functioning besides the two memory tests' scores.

2.3.2.2 Mental Health and Depression Symptom

Individual mental health condition is measured by the Centre for Epidemiologic Studies Depression Scale, or the CES-D scores. It is developed by Radloff (1977) and has become one of the most widely used tests to diagnose depression quotient. The standard CES-D measure contains 20 items and the CHARLS survey uses a simplified, 10-item CES-D measure. Specifically, respondents are asked about the frequency in which they experience a certain feeling over the week prior to the interview. The 10 items cover 8 negative feelings including feeling depressed, feeling that everything was an effort, restless sleep, feeling lonely, bothered by little things, feeling that they could not get going, feeling it hard to concentrate, feeling fearful; and 2 positive feelings of feeling happy and feeling hopeful. The 4 options of frequency are 'almost never (less than one day)', 'sometimes (1–2 days)', 'often (3–4 days)', and 'most of the time (5–7 days)', and are coded as 0, 1, 2, 3 for negative feelings and 3, 2, 1, 0 for positive feelings. The total CES-D score ranges from 0 to 30, and a lower score indicates a better mental health condition.

We also construct an indicator of depression symptoms that equals 1 if an individual scores 10 or above in the CES-D test. Previous studies have generally selected the cut-off score of 10 in the 10-item CES-D Scale which ranges from 0 to 30, and found it to have high specificity in older samples (Andresen et al., 1994; Chen and Mui, 2014). Lei et al. (2014*a*) using CHARLS to examine depressive symptoms and soio-economic status among the elderly in China also adopt the cut-off point of 10.

2.3.2.3 Physical Health Measures

Since mental health or depression is highly correlated with the number of chronic diseases, disability or limited physical functioning (Lei et al., 2014*a*), we also study retirement effects on physical frailty to explore the mechanisms through which retirement affects mental health.

Public health literature defines a retiree's physical well-being based on the absence of physical disease and functional limitation (Wang and Shi, 2014). Following the literature, we define physical frailty as an accumulated burden of chronic diseases, functional limitations, and other health-related deficits and symptoms. Specifically, we construct the physical health variable (rphyhealth) by including 12 kinds of chronic diseases (e.g., chronic lung disease), 6 kinds of limitations in activities of daily living (ADLs, e.g., dressing), and 7 kinds of limitations in mobility (e.g., walking several blocks). Physical health measure is the sum of the number of chronic diseases or limitations, with a range between 0 and 23. A higher score indicates worse physical health.

The measure can better capture the accumulated burden of physical frailty than any single one of its elements as individuals with few mobility constraints or certain types of chronic diseases may still be able to carry on with their work. The distribution of the number of ADLs is heavily skewed to the left with more than 75% reporting no limitation[6].

2.3.2.4 Labour Force Status

We define formal retirement status based on self-reported types of retirement, self-reported types of employment and self-reported types of employers. The harmonized survey published by Global Ageing Survey Data summaries the individual main labour force status as agricultural work, non-agricultural employed, non-agricultural self-

[6] There is a non-trivial amount of respondents reporting certain kinds of ADLs in one wave but no limitation at all in the following wave. The lack of persistence that usually features severe physical problems makes the number of ADLs a noisy measure of physical health if being used alone.

employed, non-agricultural unpaid family business, retired, and never worked[7][8]. If respondents engage in more than one types of jobs, they are categorized into the job type that they spend the most time on, or to non-agricultural work if they spend the same time on two activities.

There are few formal retirees receive retirement pensions going back to work after retirement and they are excluded from the sample. We create an indicator for formal retirees that equals 1 for individuals self-reporting as having completed formal retirement process and are not currently working or retired, equals 0 for individuals self-reporting as not having completed formal retirement and working in the formal sector in China, including governments, SOEs, NGOs, and private firms.

The number of years that people have spent in formal retirement is calculated as the difference between the survey year and the year when they report to complete their retirement process. It is set to 0 for those still working, and to 0.5 for people who retire in the survey year.

2.3.2.5 Control Variables

To account for factors that explain both health and labour market status, we include a large set of controls. The demographic variables contain age and age squared, indicators for educational levels, an indicator for marital status. household-level controls include the number of children, the number of household members, and the logged value of household durable assets. Finally, survey year dummies are also included to account for aggregate health shocks, time-varying reporting changes, and effects of age that are not captured by the quadratic terms. Table 2.1 shows short-forms of the dependent and explanatory variables in our empirical models and their detailed definitions.

[7] Agricultural work includes individuals doing agricultural work for their own families or others in wage for at least 10 days

[8] Respondents who have worked for at least three months during their lifetime and have searched for a new job during the last month are categorized as unemployed, otherwise as retired if they have not searched for a new job during the last month.

Table 2.1: Definitions of Covariates

rimrc	the number of words immediate recalled, ranging between 0 and 10
rdlrc	the number of words recalled 5 minutes later, ranging between 0 and 10
rmentalintact	the number of correct answers in mental intactness test, ranging between 0 and 10
rcesd10	the CES-D scores ranging between 0 and 30, a lower score indicating a better mental health condition
rdepress	an indicator of depression symptoms that equals 1 if an individual scored 10 or above in the CES-D test and equals 0 if an individual scored below 10.
rphyhealth	the sum of the number of chronic diseases or limitations ranging between 0 and 23, higher score indicating worse physical health
retire	an indicator for formal retirees that equals 1 for individuals having completed formal retirement process and are not currently working, equals 0 for those working in the formal sector
retyrs	the number of years that people have spent in formal retirement, set to 0 for those still working, and to 0.5 for people retiring in the survey year
rabove50	1 if $rage >= 50$, 0 otherwise
rabove55	1 if $rage >= 55$, 0 otherwise
rabove60	1 if $rage >= 60$, 0 otherwise
ragedif50	0 if aged 50 or below, positive and equal to the difference between individual age and 50 if above 50
ragedif55	0 if aged 55 or below, positive and equal to the difference between individual age and 55 if above 55
ragedif60	0 if aged 60 or below, positive and equal to the difference between individual age and 60 if above 60
rage	Age defined by birth year on ID, and reported age if missing ID information
rage_sq	Age square/100
rnprimary	1 if no primary school degree, including illiterate, not finishing primary school, only receiving private education (sishu), 0 otherwise (taken as reference group in the models)
rprimary	1 if highest educational attainment is finishing primary school, 0 otherwise
rsecondary	1 if highest educational attainment is finishing secondary school, 0 otherwise
rhighabove	1 if highest educational attainment is finishing high school or above (including vocational school, colleges, universities (bachelor, master or phd degree), 0 otherwise
rmarried	1 if married or partnered, 0 otherwise
hchild	number of living children
hhhres	number of people living in this household
lnhhadurbl	log (household durable assets' values)
y2013	1 if Wave 2 (2013)
y2015	1 if Wave 3 (2015)

2.3.3 Sample and Summary Statistics

We restrict the sample to individuals who are urban hukou-holders living in urban areas, are either formally retired or working in the non-agricultural sector, not miss data on health-related outcome variables that we study. The reason that we exclude from our sample non-formal retirees and agricultural workers or the self-employed is because they are not affected by the compulsory retirement policy and their retirement decisions can be confounded by many factors but not the policy age. This paper focuses instead on formal workers and using policy ages to identify their retirement probability.

The Arellano and Bond (AB) models further restrict the sample to individuals who have stayed in the survey for all three waves and not missing on the health and labour supply variables. Table 2.2 shows the missing data patterns for male and female samples that we use to estimate bivariate random effects retirement model (columns 1-4) and retirement duration model (columns 5-8) that do not account for state-dependence of health outcomes. Numbers of individuals and their corresponding percentages of the total sample are reported. It shows that over 50% of the sampled individuals stayed for at least two waves, except for the male sample used for studying effects of retirement years on cognitive functioning.

Table 2.2: Missing Patterns of the Sample

	Retirement(=1)				Retirement Years			
	Men		Women		Men		Women	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Pattern	Freq.	(%)	Freq.	(%)	Freq.	(%)	Freq.	(%)
xxx	400	22.88	366	22.92	231	19.40	246	22.22
..x	356	20.37	298	18.66	285	23.93	224	20.23
x..	287	16.42	240	15.03	194	16.29	157	14.18
.xx	225	12.87	197	12.34	158	13.27	161	14.54
xx.	195	11.16	214	13.40	122	10.24	129	11.65
.x.	180	10.30	175	10.96	125	10.50	113	10.21
x.x	105	6.01	107	6.70	76	6.38	77	6.96
Total	1748	100.00	1597	100.00	1191	100.00	1107	100.00

Notes: The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, not missing on the short-run memory test scores and the formal retirement status (columns 1-4) or years in retirement (columns 5-8).

Table 2.3 gives a summary statistics of our sample by gender and survey year. There is no clear time trend in cognitive tests' performance of men or women. Women tend to report better mental health conditions, as represented by the lower CES-D scores over time, and both men and women report slightly lower risk of depression over time. The number of physical health problems increases for both men and women over time. The ratios of formal retirees are stable over waves, but the number of years the sampled retirees have spent in retirement increases. Educational attainment, marital status, and number of household members are stable over time, and the value of household durable assets increases with time.

2.4 Identification Issues

Identifying the effect of retirement or non-working on mental health is complicated by two concerns. Firstly, individual retirement decisions and health outcomes may depend on a common set of unobserved factors. Potential unobservables include labour market attachment, personality, unobserved characteristics of pre-retirement jobs (e.g. physical and cognitive demands, work stress, workload, job satisfaction, provision of pension and health insurance), working experiences and working conditions throughout the career life, wealth and financial security, physical and mental health conditions before retirement, retirement planning, family factors including spouses' working status and health condition, events in family lives, having dependent children or parents, taking care of grandchildren or other family members, socioeconomic context include social pension and health care provision, local infrastructures, availability of care agencies and entertainment facilities etc. (Wang, 2007; Wang and Shi, 2014). The second concern is reverse causality. Individuals in worse health or well-being conditions or suffering from a decline in health conditions over time may choose to retire earlier or later than those in better health conditions.

2.4.1 Social Pension Policy as an Instrument for Retirement Decision

Labour supply of older workers is responsive to changes in retirement incentives. Evidence from the US suggests that being eligible for social pension provides strong incentives for retirement due to the higher implicit tax rates on labour income, the wealth effect of receiving benefits, while the liquidity constraint explains early retirement (French and Jones, 2012). Studying retirement incentives in the US is more complex as Medicare, a universal health insurance, also begins at age 65. Before reaching age of 65, health insurance from employers provides strong incentive for workers to continue to work (Rust and Phelan, 1997; French and Jones, 2011). French and Jones (2011) find that Medicare is as important as Social Security in predicting retirement at 65. Disability Insurance Programme that offers benefits at about 50%

Table 2.3: Summary Statistics

	Men			Women		
	2011	2013	2015	2011	2013	2015
rimrc	4.77	4.80	4.77	5.03	5.04	5.11
	(1.74)	(1.77)	(1.73)	(1.62)	(1.76)	(1.80)
rdlrc	3.70	3.84	3.75	4.14	4.16	4.07
	(1.94)	(2.00)	(2.00)	(2.11)	(2.16)	(2.05)
rmentalintact	8.31	8.34	8.31	8.05	8.13	8.03
	(2.07)	(1.95)	(2.03)	(2.21)	(2.15)	(2.20)
rcesd10	5.20	5.38	5.11	6.54	6.36	6.13
	(4.61)	(4.29)	(4.70)	(5.37)	(4.80)	(5.39)
rdepress	0.17	0.15	0.15	0.26	0.22	0.21
rphyhealth	2.67	3.15	3.45	2.88	3.49	3.78
	(2.99)	(3.30)	(3.40)	(2.89)	(3.19)	(3.24)
retire	0.61	0.61	0.59	0.76	0.76	0.73
retyrs	3.28	3.45	3.72	6.91	7.05	7.56
	(4.78)	(4.85)	(5.19)	(6.52)	(6.57)	(6.97)
rage	62.07	62.87	62.50	59.48	59.92	60.23
	(10.35)	(10.53)	(11.04)	(10.11)	(10.36)	(10.51)
rnoprimary	0.10	0.10	0.08	0.15	0.14	0.12
rprimary	0.16	0.15	0.24	0.14	0.15	0.23
rsecondary	0.27	0.27	0.25	0.30	0.29	0.28
rhighabove	0.47	0.48	0.43	0.42	0.42	0.37
rmarried	0.94	0.92	0.93	0.83	0.81	0.83
hchild	2.06	2.00	2.00	1.88	1.89	1.90
	(1.28)	(1.30)	(1.29)	(1.18)	(1.22)	(1.21)
hhhres	3.03	3.05	2.77	2.81	2.94	2.74
	(1.38)	(1.34)	(0.98)	(1.33)	(1.29)	(1.04)
lnhhadurbl	7.82	7.98	8.08	7.87	7.96	8.04
	(1.73)	(1.84)	(1.87)	(1.81)	(1.95)	(1.94)
N	1,159	1,174	1,210	1,052	1,051	1,066

Notes: The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, not missing on the short-run memory test scores and the formal retirement status.

replacement rate may induce early retirement for those evaluated as being disabled. Gruber and Wise (2005) provide a summary of public pension programmes in Europe. They tend to induce little incentive to work beyond normal retirement age due to the higher replacement rates and minimum benefit level, the actuarially unfair benefits of delaying to claim benefits[9]. In the US, unemployment benefits are paid for 1.5 years at a relatively low replacement rate, while it is more generous in European countries and provides pathways to early retirement. As for disability insurance programmes, the application is less stringent and benefits more generous in Europe than they are in the US.

Individual labour supply elasticities in response to these incentives are different, and changes over the life-cycle. Literature finds that labour supply elasticity is much higher for older workers than younger workers (French, 2005). Recent literature on the retirement effect using data from developed countries has explored the discontinuity in retirement incentives at certain ages. China provides a better opportunity to exploit the discontinuity as retirement is mandatory for urban, formal workers who reach normal retirement ages.

2.4.2 Cohort Effect

Cross-sectional studies comparing retirees with current workers may catch up with only differences between these cohorts and not the general effect of retirement on health. Cohort heterogeneity can come from cohort differences in initial conditions or mortality rates. Initial cognitive endowment, early life environment, especially differences in schooling quality, family, or origin can cause cohort heterogeneity in cognitive abilities (Mazzonna and Peracchi, 2012; Richards et al., 2004; Cunha and Heckman, 2007; Case and Paxson, 2009; Currie, 2009). Cohort-specific macroeconomic events during early life can have a lifelong effect on health, resulting in nonlinear relationships between health and age, between cognitive ageing and age. Such non-linear relationships cannot be fully controlled by age in the regression (Coe and Zamarro, 2011). Given the substantial improvements in childhood conditions and education quality, and increases in life expectancy in China in the past decades, cohort differences in initial conditions generate negative bias and over-estimate the negative retirement effect or duration effect, while cohort differences in mortality generate positive bias and underestimate the age effect (Mazzonna and Peracchi, 2012).

Given that the birth years of our sample individuals range from 1941 to 1971, the relevant macroeconomic event that can have lifelong effects on health and cognitive abilities might be the Second World War (1939-1945) and the Great Famine (1959-1961). Previous studies using cross-sectional estimators have tested for the cohort

[9]Public pension replacement rate is 80% in Spain, and is closer to 40% for the US (French and Jones, 2012)

effect by including in the model a set of cohort dummies, focusing on the cohort born during WWII (1939-1945) (Mazzonna and Peracchi, 2012; Coe and Zamarro, 2011). They find small or no significant change in health outcomes.

The Great Famine has been used as an IV for education achievement by Huang and Zhou (2013) to study the effect of education on cognition at older ages. They find that the treatment group who were born between 1948-1953, or were of primary school age during the famine, are about 14 percent less likely to finish primary school. Huang et al. (2013) find that women born during the Great Famine suffer a higher risk of mental illness, and thus provide evidence that the Great Famine may physically affect cognitive functioning through malnutrition.

If cohort heterogeneity is a fixed effect that reflects different initial conditions, Mazzonna and Peracchi (2012) suggest to difference it out by using panel dimension of the data, but this does not work if the cohort effect is due to differences in mortality, which might only be a concern for the very old sample cohort.

2.4.3 Attrition Bias

There is panel attrition bias if people in poor physical or mental health and with poor cognitive abilities are more likely to drop out of the survey or die during the survey (Behncke, 2012; Coe and Zamarro, 2011; Mazzonna and Peracchi, 2012; Atalay, Barrett and Staneva, 2019). If non-response is systematically related to health, estimates of retirement effect on health would be biased.

To give a hint on the attrition bias, we compare descriptive statistics for the attrition sample and the remaining sample. Alternatively, we can compare the estimation results using and not using individual-level, non-response adjusted weights in regression analyses. If the results differ systematically for the two models, it is likely that attrition poses a serious problem. Testing for attrition bias is not trivial and requires more work in the future. We can estimate a model for participating in the study jointly with the model we are interested in studying, and see if random effects of the two models are correlated with each other. Nonetheless, according to Mazzonna and Peracchi (2012); Atalay, Barrett and Staneva (2019), if attrition is due to fixed individual characteristics that do not change between the two waves, FD estimator remains unbiased.

2.5 Econometric Model

To investigate the overall effect of transition into retirement for older workers on outcomes of cognitive functioning, mental health condition, and subjective health for formal workers in China, we focus on estimating a bivariate random effect model with a binary endogenous treatment variable. The aim is to study separately the effect of formal retirement and the effect of years in formal retirement on cognitive outcomes, depressive symptoms, and physical well-being of retirees. Nonetheless, we recognise that health outcomes might be state-dependent and the standard OLS estimation of equation (2.1) can be inconsistent and suffer from the dynamic panel bias if the one-period lag of the dependent variable on the right-hand-side is correlated with the time-invariant, individual-specific effects (Nickell, 1981).

$$Y_{it} = \beta_0 + \beta_1 Y_{it-1} + \beta_2 Retire_{it} + \beta_3 X_{it} + \eta_{i1} + v_{it1} \tag{2.1}$$

The estimate for the lagged dependent variable contains the impact of individual-specific effects and thus is over-estimated. The FE estimator of the model would still be downward biased because Y_{it-1} is correlated with $\bar{v}_{it}$ by construction as the latter average contains v_{it-1} (Baltagi, 2008), the so-called incidental parameters problem. The most popular econometric method for fitting dynamic panel models is the GMM that uses lagged dependent variables as instruments. The method is also incorporated into the Arellano and Bond (1991) estimator (referred to as AB estimator afterwards). The AB approach provides consistent estimation for the dynamic linear model. Nonetheless, they tend to suffer from the incidental parameter bias when the time periods are small, which is the case in our study. The dynamic random effect model can distinguish between individual unobserved heterogeneity (random effects) and state dependence (the lagged dependent variable), but can suffer from initial condition problem. The Wooldridge (2005)'s approach to the initial condition only applies to non-linear models such as probit model and poisson model, while this paper focuses on linear models. Therefore, we estimate AB IV models for equation (2.1) to test for any significant state-dependence, and focus our discussion on the bivariate random-effects models in equation (2.2) that do not account for state-dependence of health outcomes. We also estimate single random-effects regressions of health outcomes on retirement variables that do not account for the endogeneity of the retirement decision for comparison.

We estimate the following random effect model for mental and physical health outcomes:

$$Y_{it} = \beta_0 + \beta_1 Retire_{it} + \beta_2 X_{it} + \eta_{i1} + v_{it1} \tag{2.2}$$

For $i = 1, ..., N$, and $t = 2, ..., T$. where X_{it} are strictly exogenous individual and household-level characteristics. η_{i1} are unobserved individual-level effects and v_{it1} are idiosyncratic errors. Y_{it} includes three sets of outcomes. Cognitive functioning is measured by the number of words being recalled immediately (IMR) that is within the range of 0 and 10, the number of words being recalled 5 minutes later (DWR) that is within the range of 0 and 10, mental intactness test scores (Intact) that is within the range of 0 and 10. Mental heath is measured by CES-D scores that is within the range of 0 and 30 and an indicator of depression symptoms that equals to 1 if CES-D scores are 10 or above. Physical health is measured by the total number of chronic diseases, limitations in daily activities (ADL), and limitations in mobility and is within the range of 0 and 23. Current retirement status, $Retire_{it}$, is a dummy variable equaling 1 for individuals who have completed retirement and are not doing any paid work at the time of survey, and equaling 0 for individuals working in the formal sector in China, including governments, SOEs, NGOs, and private firms. Equation (2.2) is estimated using linear random effects models for all health measures except for the indicator of depression symptoms, which is estimated using a random effects probit model to account for the binary nature of the dependent variable.

As $Retire_{it}$ might be endogenous to the current health condition Y_{it}, we estimate the following random effects probit model for the probability of retirement, separately for men ($Retire_{mt}$) and women ($Retire_{ft}$). We use mandatory retirement ages as instruments. The default 12 integration points are used for integration by quadrature.

$$Retire_{mt} = 1[Retire^*_{mt} = \theta_0 + \theta_1 d(Age_{mt} \geqslant 55) + \theta_2 d(Age_{mt} \geqslant 60) + \theta_3 X_{mt} + \theta_4 \bar{X_m} + \eta_{m2} + v_{mt2} > 0] \tag{2.3}$$

or for women:

$$Retire_{ft} = 1[Retire^*_{ft} = \theta_0 + \theta_1 d(Age_{ft} \geqslant 50) + \theta_2 d(Age_{ft} \geqslant 55) + \theta_3 X_{ft} + \theta_4 \bar{X_f} + \eta_{f2} + v_{ft2} > 0] \tag{2.4}$$

for $m/f = 1, ..., N^{m/f}$, and $t = 2, ..., T$, where X_{it} denotes individual and household-level controls, η_{i2} are unobserved individual-level effects and v_{it2} are idiosyncratic errors. Instruments of mandatory retirement ages for our sampled women include an indicator of passing the normal retirement age of 50 $d(Age_{ft} \geqslant 50)$ for most of the female workers, and an indicator of passing the normal retirement age of 55 $d(Age_{ft} \geqslant 55)$ for female civil servant. Similarly, an indicator of passing the early retirement age of 55, $d(Age_{mt} \geqslant 55)$, and an indicator of passing the normal retirement age of 60, $d(Age_{mt} \geqslant 60)$, are used as instruments for men' retirement probability.

We estimate a bivariate model of retirement effect on mental and physical health outcomes, and assume that the regression errors (η_{i1}, η_{i2}) share a zero-mean bivariate joint normal distribution. The model can be estimated using gsem command in Stata [10]. We report the pairwise covariance between the unobserved individual heterogeneities, $cov(\eta_{i1}, \eta_{i2})$ at the bottom of the estimation tables, in addition to the variances of η_{i1} and η_{i2}. The correlation can only be estimated from the non-missing data. A non-zero correlation between η_{i1} and η_{i2}, indicates a significant

[10] The following command is used to estimate the bivariate RE models for all health outcomes: `sem( rimrc <-retire`$_f ull$`vlist M1[id])(retire`$_f ull < -rabove60rabove55$`vlist M2[id],probit) if female==0  rrural2==0  urban`$_n bs == 1, intpoints(12)$

correlation between the time-invariant unobserved individual heterogeneities that predict transition into retirement and those predicting performance in cognitive tests, subjective or physical well-being, conditioning on the observable characteristics.

To account for the fact that mental and physical health outcomes might be state-dependent, we estimate the following AB model:

$$\Delta Y_{it} = (\alpha - 1)Y_{it-1} + \gamma_1 \Delta Retire_{it} + \gamma_2 \Delta X_{it} + \Delta u_{it} \tag{2.5}$$

where Δu_{it3} is an MA(1) process with unit root. In the case of three time periods, Y_{i1} is a valid instrument for $Y_{i2} - Y_{i1}$ as it is highly correlated with $Y_{i2} - Y_{i1}$ but not correlated with $u_{i3} - u_{i2}$ as long as u_{it} are not serially correlated. Similarly, $Retire_{i1}$ is also a valid instrument for $Retire_{i3} - Retire_{i2}$. Other instruments include ΔX_{it}, or first differences of exogenous variables. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as additional instruments.

The first differenced errors and instruments form the moment conditions. Lagged levels of the dependent variable Y_{it-2} and the endogenous retirement variable $Retire_{it-2}$, or or Y_{i1} and $Retire_{i1}$ in our case, are used as GMM-type instruments, and first differences of the strictly exogenous variables, ΔX_{it}, are used as standard instruments. Since we only have three periods of survey available, the model is exactly identified and we cannot use the Sargan test for over-identifying restrictions.

According to Blundell and Bond (1998), in the case of three time periods, for sufficiently high α or individual-specific variance of η_i, the least-squares estimate of the reduced form coefficient of $(\alpha - 1)$ can be biased upwards and be arbitrarily close to zero, as $E(Y_{i1}\eta_i) > 0$. Instrument Y_{i1} is only weakly correlated with ΔY_{i2}.

According to a limited Monte Carlo study in Arellano and Bond (1991), the estimated standard error of the two-step GMM estimator was substantially downward biased. Windmeijer (2005) suggests that the bias is caused by the weight matrix, and a more accurate approximation in finite samples can be achieved by using robust standard errors if all of the moment conditions are linear. We estimate the robust standard errors.

2.5.0.1 Retirement Duration

We re-estimate the equation (2.2) but replace the indicator of retirement with the number of years that retirees have spent in retirement, which equals to 0 for individuals still working. The aim is to study the cumulative retirement effect apart from the short-run effect of transition into retirement. Specifically, we estimate the following random effect model for mental and physical health outcomes:

$$Y_{it} = \beta_0 + \beta_1 RetYrs_{it} + \beta_2 X_{it} + \beta_3 \bar{X}_i + \eta_{i1} + v_{it1} \tag{2.6}$$

where $RetYrsit$ denotes the years that retirees have spent on retirement, calculated as the difference between the survey year and the year when they report to complete their retirement process. It is set to 0 for those still working, and to 0.5 for people who retire in the survey year.

Similarly, t account for potential endogeneity of $RetYrs_{it}$, we estimate the following random effects model, separately for men ($RetYrs_{mt}$) and women ($RetYrs_{ft}$). We use the distances to the mandatory retirement ages in years as instruments. The default 12 integration points are used for integration by quadrature.

$$RetYrs_{mt} = \theta_0 + \theta_1 Agediff55_{mt} + \theta_2 (Agediff60_{mt} + \theta_3 X_{mt} + \theta_4 \bar{X_m} + \eta_{m2} + v_{mt2} \quad (2.7)$$

or for women:

$$RetYrs_{ft} = \theta_0 + \theta_1 Agediff50_{ft} + \theta_2 (Agediff55_{ft} + \theta_3 X_{ft} + \theta_4 \bar{X_f} + \eta_{f2} + v_{ft2} \quad (2.8)$$

The instrument for $RetYrs_{m/ft}$ is the distance to the mandatory retirement ages in years. It is constructed as the time elapsed since an individual reached the normal retirement age, which is 55 or 60 ($Agediff55_{mt}$ and $Agediff60_{mt}$) for men and 50 or 55 ($Agediff50_{ft}$ and $Agediff55_{ft}$) for women, depending on the types of jobs they retire from. It is set to 0 for those below the retirement age. Other notations have the same meanings as they have in equation (2.2) to equation (2.4).

2.6 Results

Table 2.4 shows the RE estimates of the retirement probability as a function of the retirement ages (columns 1 2) and years in retirement as a function of the distance to the mandatory retirement ages in years (columns 3 4). We separate the estimation for men and women. The retirement probability model is estimated using a RE probit model and marginal effects are reported in columns 1 and 2. Column 1 shows that men are about 3.2% more likely to retire after they reach 55, which is the retirement age for blue-collar workers, though not statistically significant, and are 12.9% more likely to retire after they reach 60, which is the retirement age for white-collar workers. The fact that the retirement age of 60 has a substantially larger impact on men's probability of retirement suggests that most of the sampled retirees retire formally after 60, or retire from white-collar jobs.

Column 2 shows that the retirement age of 55 for female civil servants or managers predicts an 8.2% increase in retirement rate, while the retirement age of 50 for other white-collar female workers predicts a significantly larger rate of 14.4%. It suggests that most of the female retirees in our sample retire after 50 or from white-collar jobs other than civil service and management positions. Therefore, we expect that most of the retirement effects estimated from bivariate models come from white-collar men retiring after 60 and white-collar women retiring after 50. The fact that both indicators of retirement age are significant in predicting retirement behaviour suggests that they are valid instruments.

Column 3 shows that distance to the mandatory retirement age, constructed as the time elapsed since an individual reached the mandatory retirement age, is significant and positive in predicting the years men and women spent in formal retirement. The magnitudes are larger for men aged above 60 and women aged above 50. We also report the standard deviation of individual heterogeneity, $\sigma_{\eta_{i2}}$, and the fraction of of variance of the dependent variable due to individual heterogeneity η_{i2}, as represented by ρ. The fraction is large, justifying the use of random-effects models over pooled models. Besides the number of observations N, we also report the number of individuals as represented by N_g. On average, the same individuals are observed for more than one wave. We report the overall R square ($r2\_o$) of single linear RE models (columns 3 and 4) and log likelihood (ll) of RE probit models and bivariate RE models in the following tables.

2.6.1 Effect of Transition into Retirement

Tables 2.5 shows the bivariate RE estimation of the retirement effect on men and women's cognitive functioning. The retirement probability is estimated using a RE probit model and cognitive functioning is estimated using a linear RE model. Recall that we study three measures of cognitive functioning, the number of words being immediately recalled (IMR), the number of words being recalled 5 minutes later (DMR), and the number of tasks being correctly performed in the mental intactness test (Intact), separately for men and women. The retirement indicator captures the short-run effect of transition into retirement. Random effects of the two models are allowed to be correlated with each other to account for unobserved, time-invariant individual heterogeneity that determines both retirement decisions and health outcomes, which can be tested for by looking at the significant levels of the covariance of η_{i1} and η_{i2}. We report at the bottom of the table both the number of observations, N, variances of the time-invariant individual-specific terms η_{i1} in equation (2.2) and η_{i2} in equation (2.3) and equation (2.4), covariance of η_{i1} and η_{i2} that indicates the endogeneity of retirement decision.

As shown in Table 2.5, variances of the random effects are significant for both the equation of cognitive functioning ($var(\eta_{i1})$) and the equation of retirement status ($var(\eta_{i2})$), justifying the use of RE models over pooled models. Cross-equation covariances of the two random effects, $cov(\eta_{i1}, \eta_{i2})$, suggest that there exist unobserved, time-invariant individual heterogeneities that predict both men's transition into retirement and their lower scores in cognitive tests. Not accounting for the endogeneity of men's retirement would cause a downward bias in estimating

Table 2.4: Determinants of Retirement Probability and Retirement Duration

	Retire		RetYrs	
	(1)	(2)	(3)	(4)
	Men	Women	Men	Women
rabove50		0.144***		
		(0.019)		
rabove55	0.032	0.082***		
	(0.021)	(0.018)		
rabove60	0.129***			
	(0.016)			
ragedif50				0.749***
				(0.038)
ragedif55			0.328***	0.298***
			(0.055)	(0.037)
ragedif60			0.427***	
			(0.054)	
rage	0.009	0.028***	-0.534*	0.665***
	(0.013)	(0.008)	(0.293)	(0.120)
rage_sq	0.008	-0.013**	0.573**	-0.664***
	(0.010)	(0.007)	(0.292)	(0.126)
rprimary	0.017	0.022	0.215	0.735*
	(0.027)	(0.027)	(0.417)	(0.425)
rsecondary	0.044*	0.110***	-0.060	1.501***
	(0.025)	(0.025)	(0.403)	(0.397)
rhighabove	0.014	0.037	-0.677*	0.222
	(0.025)	(0.025)	(0.394)	(0.393)
rmarried	-0.043	-0.025	0.049	-0.127
	(0.057)	(0.052)	(0.211)	(0.129)
hchild	0.027*	0.011	0.059	-0.037
	(0.015)	(0.013)	(0.068)	(0.039)
hhhres	-0.004	0.009	0.005	0.002
	(0.006)	(0.007)	(0.027)	(0.020)
lnhhadurbl	0.001	-0.004	-0.036**	-0.008
	(0.004)	(0.004)	(0.017)	(0.012)
y2013	-0.015	-0.011	0.026	0.193***
	(0.009)	(0.011)	(0.050)	(0.043)
y2015	-0.030***	-0.030**	0.026	0.352***
	(0.010)	(0.012)	(0.068)	(0.066)
_cons			14.917**	-16.618***
			(7.410)	(3.116)
$\sigma_{\eta_{i2}}$	2.073	2.340	2.967	3.386
σ_v			0.684	0.501
ρ	0.811	0.846	0.950	0.979
Obervation	3445	3086	2239	2109
Individuals	1917	1693	1318	1155
R^2			0.581	0.721
log likelihood	-856.3	-817.2		

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

Table 2.5: Bivariate RE Estimation of Retirement Effect on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.378***	0.026	0.502***	-0.017	0.312*	-0.523***
	(0.144)	(0.155)	(0.164)	(0.190)	(0.177)	(0.202)
rage	0.027	0.039	-0.019	-0.016	-0.012	0.137***
	(0.040)	(0.038)	(0.045)	(0.047)	(0.047)	(0.050)
rage_sq	-0.062**	-0.052*	-0.040	-0.019	-0.018	-0.120***
	(0.030)	(0.030)	(0.034)	(0.036)	(0.036)	(0.039)
rprimary	0.520***	0.656***	0.690***	0.833***	0.553***	1.182***
	(0.138)	(0.122)	(0.154)	(0.151)	(0.163)	(0.164)
rsecondary	0.777***	0.961***	0.935***	1.170***	0.936***	1.831***
	(0.134)	(0.118)	(0.149)	(0.146)	(0.159)	(0.158)
rhighabove	1.064***	1.436***	1.091***	1.653***	1.213***	2.122***
	(0.129)	(0.115)	(0.144)	(0.143)	(0.153)	(0.154)
rmarried	-0.295	0.086	-0.289	0.180	0.267	0.071
	(0.266)	(0.251)	(0.313)	(0.302)	(0.319)	(0.301)
hchild	-0.157**	-0.020	-0.138	-0.098	0.053	0.035
	(0.077)	(0.068)	(0.090)	(0.081)	(0.093)	(0.084)
hhhres	-0.062	0.002	-0.002	-0.034	0.014	0.033
	(0.040)	(0.044)	(0.047)	(0.053)	(0.050)	(0.054)
lnhhadurbl	0.045*	0.056**	0.028	0.072***	0.062**	0.052*
	(0.023)	(0.023)	(0.027)	(0.027)	(0.029)	(0.028)
y2013	0.072	0.030	0.214***	0.051	0.004	0.103
	(0.064)	(0.065)	(0.074)	(0.078)	(0.077)	(0.080)
y2015	-0.026	0.084	0.040	-0.067	-0.074	-0.060
	(0.066)	(0.068)	(0.077)	(0.082)	(0.080)	(0.084)
η_{i1}	1.000	1.000	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)	(.)	(.)
_cons	3.669***	3.179***	4.168***	4.235***	6.905***	2.617*
	(1.280)	(1.154)	(1.446)	(1.414)	(1.525)	(1.523)
var(η_{i1})	0.838***	0.630***	0.904***	1.070***	1.051***	1.509***
	(0.074)	(0.069)	(0.093)	(0.106)	(0.106)	(0.125)
var(η_{i2})	4.606***	6.588***	4.489***	6.539***	4.578***	6.595***
	(1.037)	(1.711)	(0.985)	(1.696)	(1.009)	(1.721)
cov(η_{i1}, η_{i2})	-0.356*	-0.327	-0.669***	-0.429	-0.591**	0.371
	(0.196)	(0.224)	(0.219)	(0.277)	(0.243)	(0.308)
Observation	3445	3086	3445	3086	3445	3086
log likelihood	-6503	-5966	-6913	-6525	-7144	-6731

Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

retirement effects on men's cognitive functioning, as shown in Table 2.6.

Estimation of the retirement effect in Table 2.5 shows that transition into formal retirement significantly improves men's scores in both memory tests including IMR and DWR (columns 1 and 3), and in the mental intactness test (column 5). The retirement effect is not significant on women's scores in memory tests (columns 2 and 4), and becomes significantly negative on their scores in the mental intactness test (column 6).

Compared with the AB IV estimation in Table B.2 in Appendix B, we can see that the signs of retirement are consistent for men, but the retirement effect is only significant and positive in predicting men's episodic memory (column 3). For women, the AB IV estimations of the retirement effects are negative but insignificant across all three cognitive test outcomes. Apart from not accounting for state-dependence, the RE estimation also differs from the AB estimations in the selected sample as the AB models select on individuals staying for all three waves and have substantially smaller sample sizes, and may face with attrition bias. State-dependence is not significant across all AB models of cognitive functioning, therefore we focus on models that do not account for state-dependence.

Estimates of the other covariates in Table 2.5 show that for both men and women, higher educational levels are related to higher scores in cognitive tests, including both memory tests and the mental intactness test. The educational gradient is stronger for women than for men. Higher levels of household assets values, as measured by the logged values of the household durable assets, also predict better performance in cognitive functioning tests for both men and women.

In Table 2.6, we re-estimate the models in Table 2.5 using single RE models that do not account for the potential endogeneity of retirement decisions, and study the correlation between retirement and cognitive functioning, separately for men and women. Compared to the estimated positive retirement effects on men's cognitive outcomes in Table 2.5, the magnitudes of the effects drop substantially and become insignificant. Retirement effects also decline in terms of magnitudes and statistical significance in predicting women's scores in the mental intactness test (column 6), and become significant in predicting their scores in the memory tests (columns 2 and 4). The RE estimations are biased because they do not account for the endogeneity of formal retirement. Individuals suffering from cognitive declines may choose to retire earlier before the formal retirement ages. Correlations between unobserved, time-invariant individual heterogeneities ($cov(\eta_{i1}, \eta_{i2})$) also suggest that male retirees overall perform worse than non-retirees in cognitive functioning tests. Not accounting for the unobserved effects would downwardly bias the positive retirement effect on cognitive functioning of men in our sample, who tend to take white-collar jobs and retire after 60.

In Table 2.7, we study the retirement effect on subjective and physical well-being

Table 2.6: RE Estimation of Retirement Effect on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.112	-0.307**	0.084	-0.410***	-0.036	-0.493***
	(0.113)	(0.126)	(0.130)	(0.152)	(0.135)	(0.162)
rage	0.090**	0.091*	0.060	0.058	0.057	0.125**
	(0.043)	(0.050)	(0.045)	(0.061)	(0.049)	(0.062)
rage_sq	-0.001***	-0.001**	-0.001***	-0.001	-0.001*	-0.001**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
rprimary	0.576***	0.562***	0.703***	0.700***	0.594***	1.056***
	(0.153)	(0.131)	(0.172)	(0.163)	(0.211)	(0.215)
rsecondary	0.746***	0.912***	0.876***	1.086***	0.863***	1.705***
	(0.152)	(0.129)	(0.170)	(0.159)	(0.200)	(0.201)
rhighabove	1.068***	1.389***	1.055***	1.563***	1.104***	2.025***
	(0.149)	(0.128)	(0.166)	(0.157)	(0.194)	(0.199)
rmarried	-0.328	0.014	-0.315	0.157	0.284	-0.083
	(0.278)	(0.263)	(0.283)	(0.312)	(0.348)	(0.299)
hchild	-0.172**	-0.038	-0.170*	-0.141	0.069	0.044
	(0.073)	(0.054)	(0.089)	(0.099)	(0.114)	(0.083)
hhhres	-0.067	0.001	-0.025	-0.024	0.010	0.050
	(0.041)	(0.045)	(0.048)	(0.057)	(0.049)	(0.059)
lnhhadurbl	0.052*	0.050**	0.034	0.081***	0.067*	0.048*
	(0.028)	(0.023)	(0.032)	(0.026)	(0.035)	(0.029)
y2013	0.064	0.069	0.220***	0.099	0.008	0.127
	(0.067)	(0.068)	(0.079)	(0.083)	(0.079)	(0.083)
y2015	-0.003	0.110	0.073	-0.007	-0.080	-0.042
	(0.070)	(0.071)	(0.082)	(0.086)	(0.088)	(0.089)
_cons	1.850	2.046	1.493	2.397	4.888***	3.410*
	(1.377)	(1.452)	(1.467)	(1.798)	(1.534)	(1.833)
Observation	2836	2603	2834	2603	2868	2629
Individuals	1633	1472	1629	1470	1651	1488
R^2	0.194	0.244	0.194	0.229	0.121	0.184
$\sigma_{\eta_{i1}}$	0.880	0.760	0.877	0.987	1.011	1.239
$\sigma_{v_{it1}}$	1.312	1.304	1.558	1.567	1.623	1.583
ρ	0.311	0.254	0.241	0.284	0.280	0.380

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

for men and women using bivariate RE models. Subjective well-being is measured by the CES-D test score that is within the range of 0 and 30, and the lower scores indicate the better mental health conditions. The indicator of depression symptoms is constructed based on whether the individual scores 10 or above in the CES-D test. Apart from subjective well-being, we also study physical health problems as measured by the total number of chronic diseases, functional limitations, and mobility limitations. Variances of the random effects are statistically significant for both the models of subjective or physical well-being ($var(\eta_{i1})$) and the equation of retirement status ($var(\eta_{i2})$), justifying the use of RE models instead of pooled models. Cross-equation covariances of the two random effects, $cov(\eta_{i1}, \eta_{i2})$, suggest that there exist unobserved, time-invariant individual heterogeneities that predict both transition into retirement and more physical health problems for men and women. The correlation between unobserved effects that predict retirement and worse subjective well-being is only marginally significant. Not accounting for the endogeneity of retirement would cause an upward bias in estimating retirement effects on men's and women's physical problems, as shown in Table 2.8.

The bivariate RE estimations suggest that transitions into retirement do not significantly predict the subjective well-being of either men or women and physical problems of men, but reduce the number of physical health conditions for women. Subjective health measures can be subject to role bias and change systematically with labour force statuses. Retirees can feel healthier if their daily activities are less physically demanding than they were at work, and thus they feel fewer physical limitations (Neuman, 2008).

The estimates for other covariates show that having achieved high-school or above education significantly improves the mental health status of both men and women, while being married and living with partners, and higher values of household durable assets reduce the depression symptoms and depression probability only for men. There is a time trend of more reported physical problems for both men and women, and lower depression risk only for women, as suggested by the survey year dummies.

Again, in Table 2.8, we re-estimate the models in Table 2.7 but do not account for the potential endogeneity of retirement decisions. We study only the correlation between retirement and mental or physical well-being, separately for men and women. Compared to the statistically insignificant retirement effects in Table 2.7, the magnitudes and significant levels of retirement effects increase for both men and women. Both male and female retirees report worse subjective mental health status than non-retirees do, though the differences are only marginally significant. Retirees also report significantly more physical health problems than non-retirees do, especially among women. The retirement effects are more significant and of larger magnitudes when we do not account for its endogeneity. This suggests that retirees overall report worse mental health status and more physical health problems than non-retirees do,

Table 2.7: Bivariate RE Estimation of Retirement Effect on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.468	-0.175	0.192	0.148	-0.227	-0.517**
	(0.343)	(0.504)	(0.202)	(0.183)	(0.187)	(0.259)
rage	0.090	0.217*	0.017	0.035	-0.249***	0.193***
	(0.108)	(0.125)	(0.059)	(0.047)	(0.068)	(0.070)
rage_sq	-0.079	-0.167*	-0.018	-0.031	0.276***	-0.066
	(0.083)	(0.098)	(0.044)	(0.037)	(0.052)	(0.055)
rprimary	-0.293	-0.129	-0.106	0.087	0.065	0.123
	(0.393)	(0.419)	(0.186)	(0.148)	(0.272)	(0.261)
rsecondary	-0.982**	-0.693*	-0.276	-0.198	-0.377	0.132
	(0.383)	(0.404)	(0.183)	(0.144)	(0.261)	(0.245)
rhighabove	-1.051***	-1.081***	-0.421**	-0.356**	-0.527**	-0.140
	(0.369)	(0.394)	(0.178)	(0.143)	(0.252)	(0.241)
rmarried	-2.073***	-1.083	-0.983***	0.041	-0.106	-0.348
	(0.636)	(0.730)	(0.372)	(0.296)	(0.291)	(0.279)
hchild	0.045	0.109	0.243**	0.087	0.045	0.006
	(0.181)	(0.199)	(0.115)	(0.087)	(0.080)	(0.076)
hhhres	-0.021	-0.105	-0.022	-0.027	-0.110**	-0.005
	(0.099)	(0.130)	(0.059)	(0.053)	(0.045)	(0.049)
lnhhadurbl	-0.180***	-0.044	-0.096***	-0.021	-0.023	0.020
	(0.056)	(0.066)	(0.033)	(0.027)	(0.025)	(0.025)
y2013	0.225	-0.179	-0.094	-0.165**	0.399***	0.450***
	(0.155)	(0.193)	(0.092)	(0.080)	(0.070)	(0.074)
y2015	-0.013	-0.326	-0.146	-0.214**	0.741***	0.746***
	(0.164)	(0.204)	(0.097)	(0.085)	(0.081)	(0.086)
η_{i1}	1.000	1.000	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)	(.)	(.)
_cons	6.822**	3.065	-0.577	-1.197	8.403***	-5.846***
	(3.449)	(3.796)	(1.886)	(1.410)	(2.206)	(2.150)
var(η_{i1})	9.627***	11.270***	1.377***	0.968***	6.435***	6.236***
	(0.574)	(0.796)	(0.250)	(0.166)	(0.270)	(0.291)
var(η_{i2})	4.699***	6.572***	4.704***	6.640***	4.957***	6.905***
	(1.064)	(1.700)	(1.060)	(1.739)	(1.081)	(1.677)
cov(η_{i1}, η_{i2})	0.631	1.384*	0.330	0.231	2.351***	3.576***
	(0.491)	(0.794)	(0.262)	(0.259)	(0.477)	(0.732)
Observation	3445	3086	3445	3086	3445	3086
log likelihood	-9660	-9441	-2084	-2265	-8169	-7328

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

likely due to the fact that individuals suffering from worse mental and physical health conditions choose to retire before the formal retirement ages. Not accounting for the endogeneity of retirement would upwardly bias the negative retirement effect on the mental and physical health of men and women in our sample.

Table 2.8: RE Estimation of Retirement Effect on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.860**	0.767*	0.060**	0.056*	0.535***	0.746***
	(0.347)	(0.402)	(0.026)	(0.032)	(0.169)	(0.167)
rage	0.039	0.177	-0.000	0.017	-0.387***	0.018
	(0.113)	(0.134)	(0.009)	(0.011)	(0.075)	(0.084)
rage_sq	-0.000	-0.001	-0.000	-0.000	0.004***	0.001
	(0.001)	(0.001)	(0.000)	(0.000)	(0.001)	(0.001)
rprimary	-0.181	-0.220	-0.003	0.022	0.181	0.088
	(0.454)	(0.464)	(0.037)	(0.037)	(0.346)	(0.290)
rsecondary	-0.721	-0.774*	-0.020	-0.049	-0.271	-0.101
	(0.440)	(0.426)	(0.035)	(0.034)	(0.321)	(0.271)
rhighabove	-0.816*	-1.132***	-0.045	-0.077**	-0.417	-0.372
	(0.433)	(0.423)	(0.034)	(0.034)	(0.312)	(0.261)
rmarried	-2.018***	-0.598	-0.156**	0.039	-0.132	-0.170
	(0.736)	(0.928)	(0.066)	(0.067)	(0.406)	(0.253)
hchild	0.014	0.200	0.034**	0.034**	0.050	-0.033
	(0.199)	(0.176)	(0.016)	(0.017)	(0.111)	(0.070)
hhhres	-0.025	-0.146	-0.005	-0.008	-0.105*	-0.023
	(0.099)	(0.151)	(0.009)	(0.013)	(0.055)	(0.060)
lnhhadurbl	-0.187***	-0.029	-0.017***	-0.005	-0.020	0.020
	(0.059)	(0.069)	(0.006)	(0.006)	(0.035)	(0.028)
y2013	0.205	-0.263	-0.017	-0.044**	0.453***	0.479***
	(0.164)	(0.201)	(0.015)	(0.018)	(0.077)	(0.078)
y2015	-0.023	-0.397*	-0.022	-0.055***	0.802***	0.802***
	(0.181)	(0.219)	(0.016)	(0.019)	(0.092)	(0.094)
_cons	8.126**	3.583	0.372	-0.121	12.865***	-0.274
	(3.656)	(4.073)	(0.294)	(0.335)	(2.372)	(2.510)
Observation	2895	2664	2895	2664	3014	2675
Individuals	1659	1498	1659	1498	1643	1437
R^2	0.076	0.063	0.055	0.056	0.193	0.204
$\sigma_{\eta_{i1}}$	3.080	3.230	0.210	0.220	2.484	2.426
$\sigma_{v_{it1}}$	3.176	3.875	0.292	0.347	1.448	1.447
ρ	0.485	0.410	0.341	0.287	0.746	0.738

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

2.6.2 Effects of Years in Retirement

In this section, we replace the retirement indicator with the number of years that individuals have spent in retirement to study the cumulative retirement effect over time, in addition to the short-run effect of transition into retirement. The sample sizes drop as some retirees missing the year when they retired, which is used to construct the years they have spent in retirement.

As shown in Table 2.9, variances of the random effects are significant for both the equation of cognitive functioning ($var(\eta_{i1})$) and the equation of retirement years ($var(\eta_{i2})$), justifying the use of RE models instead of pooled models. Cross-equation covariance of the two random effects, $cov(\eta_{i1}, \eta_{i2})$ are only significant in determining women's scores in the mental intactness test, suggesting that there exist unobserved time-invariant individual effects that are predict both more years in retirement and lower test scores for women. The unobserved effects can be related to cognitive ageing or the loss of intellectually stimulating activities and social activities in the working environment that accelerate female retirees' cognitive declines and cause them to perform worse in mental intactness tests. Not accounting for the endogeneity of women's years in retirement would cause a downward bias in estimating the effect of retirement years on men's and women's cognitive declines, as shown in Table 2.10.

Compared to the bivariate RE estimation of the retirement effects on men's cognitive functioning in Table 2.5, the effects of years in retirement in Table 2.9 decline substantially and become not statistically significant across all models except for the model of women's mental intactness scores. It means that the positive short-run effects of transition into retirement disappear in the long-term after male retirees spend more time in retirement. The cumulative effect of retirement duration is significant and positive for women's mental intactness scores, in contrast to the negative and significant effect of the transition into retirement (column 6 in Table 2.5). It suggests that the negative short-run effects of transition into retirement disappear after women spend more time in retirement. This can be due to the fact that women retire much earlier at 50 compared to men in China and women in other developed countries do. They might take up other mentally stimulating activities after retirement that compensate for the loss of stimulation and human capital investment in the working environment, which may have started to decline before they formally retired. Covariates show similar effects compared to those in Table 2.5.

For comparison, we also report RE estimation of the effects of retirement and retirement duration without accounting for their endogeneity in Table 2.10. The effect of retirement duration declines in magnitude in models not accounting for its potential endogeneity, similarly for men and women. The negative correlation between retirement duration and cognitive functioning suggests that retirees spending more time in retirement overall experience higher levels of cognitive declines than non-retirees or retirees spending fewer years in retirement do. It can be due to the cognitive

Table 2.9: Bivariate RE Estimation of Years in Retirement on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.069	0.038	-0.006	0.026	-0.044	0.182***
	(0.059)	(0.039)	(0.070)	(0.049)	(0.069)	(0.048)
rage	0.170	-0.073	-0.141	-0.147*	-0.278	-0.116
	(0.200)	(0.066)	(0.236)	(0.081)	(0.231)	(0.082)
rage_sq	-0.198	0.029	0.095	0.086	0.247	-0.024
	(0.197)	(0.071)	(0.233)	(0.088)	(0.228)	(0.089)
rprimary	0.721***	0.581***	0.985***	0.674***	0.759***	0.735***
	(0.198)	(0.168)	(0.218)	(0.208)	(0.224)	(0.226)
rsecondary	0.928***	0.861***	1.164***	0.999***	0.987***	1.350***
	(0.189)	(0.160)	(0.208)	(0.199)	(0.214)	(0.215)
rhighabove	1.288***	1.367***	1.332***	1.547***	1.229***	1.895***
	(0.190)	(0.149)	(0.210)	(0.185)	(0.216)	(0.201)
rmarried	-0.543	0.089	-0.694	0.181	0.839*	0.486
	(0.378)	(0.298)	(0.446)	(0.367)	(0.433)	(0.357)
hchild	-0.126	-0.096	-0.078	-0.177	0.119	0.025
	(0.121)	(0.092)	(0.142)	(0.113)	(0.138)	(0.112)
hhhres	-0.076	-0.002	-0.054	-0.038	-0.034	0.018
	(0.048)	(0.049)	(0.057)	(0.060)	(0.055)	(0.059)
lnhhadurbl	0.065**	0.100***	0.060	0.113***	0.015	0.053
	(0.032)	(0.029)	(0.038)	(0.036)	(0.037)	(0.036)
y2013	0.098	0.085	0.290***	0.050	-0.067	0.072
	(0.080)	(0.078)	(0.094)	(0.096)	(0.091)	(0.095)
y2015	0.112	0.088	0.219**	-0.105	-0.074	-0.129
	(0.082)	(0.082)	(0.096)	(0.102)	(0.093)	(0.101)
η_{i1}	1.000	1.000	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)	(.)	(.)
_cons	0.141	6.730***	7.223	8.403***	13.634**	12.557***
	(5.225)	(1.733)	(6.156)	(2.152)	(6.027)	(2.187)
var(η_{i1})	0.773***	0.664***	0.728***	1.032***	0.885***	1.597***
	(0.128)	(0.093)	(0.113)	(0.134)	(0.114)	(0.289)
var(η_{i2})	8.809***	12.500***	8.809***	12.498***	8.808***	12.474***
	(0.355)	(0.540)	(0.355)	(0.540)	(0.355)	(0.538)
cov(η_{i1}, η_{i2})	-0.791	-0.579	-0.241	-0.454	0.194	-2.585***
	(0.551)	(0.539)	(0.651)	(0.666)	(0.640)	(0.668)
var(v_{it1})	1.735***	1.670***	2.463***	2.523***	2.296***	2.489***
	(0.081)	(0.078)	(0.114)	(0.119)	(0.108)	(0.116)
var(v_{it2})	0.461***	0.258***	0.461***	0.258***	0.462***	0.258***
	(0.021)	(0.012)	(0.021)	(0.012)	(0.021)	(0.012)
Observation	2239	2109	2239	2109	2239	2109
log likelihood	-8235	-7643	-8514	-8055	-8533	-8098

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The distance to the mandatory retirement ages in years, constructed as the time elapsed since an individual reached the normal retirement age, which is 55 or 60 ($Agediff55_{it}$ and $Agediff60_{it}$) for men and 50 or 55 ($Agediff50_{it}$ and $Agediff55_{it}$) for women, are used as instruments for years in retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

ageing process that comes with age instead of retirement, as retirees who report more years in retirement tend to be older than those spending fewer years in retirement and non-retirees. Not accounting for the unobserved effects would downwardly bias the effect of years in retirement on cognitive functioning of retirees.

Bivariate RE estimates in Table 2.11 suggest that the effect of retirement duration is insignificant on the subjective and physical health status of either men or women. This is similar to the null pattern of retirement effects on their mental or physical health (Tables 2.7). Random effects of the health model or the retirement duration model are significant, as suggested by $var(\eta_{i1})$ and $var(\eta_{i1})$. Cross-equation covariances of the two random-effects, $cov(\eta_{i1}, \eta_{i2})$, suggest that there is no unobserved time-invariant individual effect that significantly determines both subjective or physical well-being and the years people staying in retirement.

RE estimation of the correlation between health and retirement variables in Table 2.12 suggests that retirees who have stayed for longer in retirement tend to report more physical problems than non-retirees or retirees who have stayed for fewer years in retirement do, after controlling for age effects. Not accounting for the endogeneity of the timing of retirement can cause an upward bias on the effect of retirement duration on individual physical health condition.

In summary, we find that retirement has a short-run 'honey moon' effect on men's cognitive functioning that disappear with the number of years they spend in retirement. Transition into retirement has a short-run negative effect on women's cognitive functioning that reverses over time when they spend more years not providing a formal labour supply. There is no significant change in formal retirees' mental or physical health status shortly after their retirement or coming with the years they spend in retirement. The gender-difference in the retirement effects on cognitive functioning can be explained by the fact that there is a 10-year gap in the normal retirement ages for most of the male and female retirees in our sample. The social activities they take after retirement also matter, as activities that can compensate for the loss of intellectual stimulation and social network in the working environment are likely to slow down the cognitive decline process of retirees. Female retirees are more likely to take informal labour such as caring for grandchildren after retirement.

2.7 The Underlying Mechanisms between Mental Health and Retirement

To study mechanisms underlying the retirement effect on mental and physical outcomes, we look at changes in the probability of taking up various types of social activities after retirement. We estimate bivariate probit random effects models for the effect of transition into retirement on individual participation in 9 types of activities

Table 2.10: RE Estimation of Years in Retirement on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	-0.015	-0.002	-0.032**	-0.006	-0.024	-0.001
	(0.012)	(0.011)	(0.013)	(0.013)	(0.016)	(0.017)
rage	-0.084	-0.104	-0.220**	-0.170**	-0.216**	-0.258**
	(0.098)	(0.065)	(0.109)	(0.072)	(0.104)	(0.102)
rage_sq	0.059	0.081	0.174*	0.126**	0.184*	0.213**
	(0.088)	(0.060)	(0.098)	(0.063)	(0.094)	(0.096)
rprimary	0.731***	0.623***	0.986***	0.711***	0.756***	0.918***
	(0.208)	(0.182)	(0.250)	(0.227)	(0.284)	(0.275)
rsecondary	0.923***	0.926***	1.161***	1.053***	0.989***	1.639***
	(0.196)	(0.164)	(0.239)	(0.205)	(0.263)	(0.248)
rhighabove	1.235***	1.366***	1.315***	1.549***	1.243***	1.891***
	(0.196)	(0.165)	(0.239)	(0.206)	(0.261)	(0.247)
rmarried	-0.556	0.078	-0.705*	0.170	0.840**	0.455
	(0.404)	(0.298)	(0.386)	(0.383)	(0.429)	(0.311)
hchild	-0.119	-0.102	-0.078	-0.179	0.117	-0.000
	(0.118)	(0.074)	(0.141)	(0.162)	(0.165)	(0.089)
hhhres	-0.077	-0.005	-0.054	-0.041	-0.033	0.007
	(0.048)	(0.050)	(0.058)	(0.065)	(0.054)	(0.064)
lnhhadurbl	0.062*	0.100***	0.060	0.113***	0.016	0.051
	(0.036)	(0.026)	(0.046)	(0.033)	(0.047)	(0.035)
y2013	0.103	0.093	0.291***	0.057	-0.068	0.107
	(0.083)	(0.079)	(0.095)	(0.099)	(0.089)	(0.095)
y2015	0.119	0.108	0.222**	-0.090	-0.074	-0.042
	(0.082)	(0.080)	(0.098)	(0.098)	(0.097)	(0.100)
_cons	6.601**	7.037***	9.228***	8.609***	12.065***	14.027***
	(2.791)	(1.861)	(3.122)	(2.131)	(2.931)	(2.791)
Observation	1993	1950	1992	1949	2009	1970
Individuals	1182	1098	1179	1097	1195	1104
R^2	0.086	0.145	0.095	0.131	0.063	0.110
$\sigma_{\eta_{i1}}$	0.809	0.818	0.820	1.042	0.928	1.089
$\sigma_{v_{it1}}$	1.328	1.283	1.577	1.577	1.519	1.568
ρ	0.271	0.289	0.213	0.304	0.272	0.325

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

Table 2.11: Bivariate RE Estimation of Years in Retirement on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.178	-0.106	-0.019	-0.020	-0.023	0.030
	(0.141)	(0.127)	(0.084)	(0.067)	(0.054)	(0.057)
rage	0.562	0.323	-0.222	0.161	-0.158	0.225**
	(0.487)	(0.217)	(0.289)	(0.125)	(0.203)	(0.110)
rage_sq	-0.535	-0.201	0.214	-0.119	0.211	-0.120
	(0.477)	(0.234)	(0.284)	(0.141)	(0.195)	(0.113)
rprimary	-0.815	-0.230	-0.312	0.110	-0.067	0.051
	(0.579)	(0.574)	(0.276)	(0.216)	(0.334)	(0.319)
rsecondary	-0.847	-0.778	-0.278	-0.179	-0.277	-0.158
	(0.556)	(0.546)	(0.262)	(0.216)	(0.312)	(0.291)
rhighabove	-1.153**	-0.963*	-0.585**	-0.205	-0.378	-0.094
	(0.554)	(0.510)	(0.268)	(0.192)	(0.308)	(0.278)
rmarried	-3.514***	-0.632	-0.972*	0.004	-0.406	-0.326
	(0.909)	(0.878)	(0.514)	(0.367)	(0.356)	(0.330)
hchild	0.032	0.174	0.316*	0.175	0.017	-0.081
	(0.292)	(0.269)	(0.185)	(0.125)	(0.113)	(0.100)
hhhres	-0.005	-0.120	-0.034	-0.013	-0.085*	-0.024
	(0.116)	(0.144)	(0.071)	(0.060)	(0.044)	(0.053)
lnhhadurbl	-0.251***	-0.035	-0.128***	-0.017	-0.006	-0.017
	(0.077)	(0.086)	(0.048)	(0.036)	(0.028)	(0.032)
y2013	0.582***	-0.154	0.032	-0.140	0.320***	0.378***
	(0.194)	(0.234)	(0.119)	(0.101)	(0.075)	(0.088)
y2015	0.104	-0.340	-0.105	-0.274**	0.536***	0.686***
	(0.203)	(0.250)	(0.125)	(0.109)	(0.086)	(0.102)
η_{i1}	1.000	1.000	1.000	1.000	1.000	1.000
	(.)	(.)	(.)	(.)	(.)	(.)
_cons	-5.092	-0.719	6.045	-5.065*	5.045	-5.904*
	(12.810)	(5.763)	(7.576)	(3.020)	(5.454)	(3.057)
var(η_{i1})	9.755***	11.302***	1.369***	1.082***	4.471***	4.709***
	(0.767)	(1.001)	(0.332)	(0.217)	(0.241)	(0.268)
var(η_{i2})	8.807***	12.492***	8.835***	12.493***	8.808***	12.496***
	(0.355)	(0.539)	(0.356)	(0.539)	(0.355)	(0.540)
cov(η_{i1}, η_{i2})	-1.223	1.608	0.255	0.245	0.788	0.506
	(1.320)	(1.750)	(0.787)	(0.879)	(0.524)	(0.797)
var(v$_{it1}$)	9.448***	14.361***			1.343***	1.914***
	(0.450)	(0.664)			(0.063)	(0.088)
var(v$_{it2}$)	0.462***	0.258***	0.465***	0.258***	0.462***	0.258***
	(0.021)	(0.012)	(0.022)	(0.012)	(0.021)	(0.012)
Observation	2239	2109	2239	2109	2239	2109
log likelihood	-10267	-10028	-5342	-5074	-8855	-8386

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The distance to the mandatory retirement ages in years, constructed as the time elapsed since an individual reached the normal retirement age, which is 55 or 60 ($Agediff55_{it}$ and $Agediff60_{it}$) for men and 50 or 55 ($Agediff50_{it}$ and $Agediff55_{it}$) for women, are used as instruments for years in retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

Table 2.12: RE Estimation of Years in Retirement on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.051	0.007	0.001	-0.001	0.052**	0.064***
	(0.040)	(0.035)	(0.003)	(0.003)	(0.026)	(0.020)
rage	0.182	0.412***	-0.021	0.028**	0.071	0.253***
	(0.284)	(0.157)	(0.022)	(0.013)	(0.149)	(0.084)
rage_sq	-0.150	-0.347***	0.019	-0.022**	-0.021	-0.166**
	(0.260)	(0.134)	(0.020)	(0.011)	(0.138)	(0.075)
rprimary	-0.805	-0.328	-0.067	0.021	-0.074	0.030
	(0.685)	(0.625)	(0.053)	(0.049)	(0.392)	(0.353)
rsecondary	-0.862	-0.948*	-0.066	-0.054	-0.270	-0.208
	(0.661)	(0.536)	(0.051)	(0.042)	(0.373)	(0.291)
rhighabove	-1.238*	-0.952*	-0.104**	-0.046	-0.329	-0.091
	(0.653)	(0.545)	(0.050)	(0.043)	(0.366)	(0.287)
rmarried	-3.498***	-0.609	-0.190*	0.011	-0.403	-0.323
	(1.190)	(1.179)	(0.098)	(0.078)	(0.323)	(0.293)
hchild	0.039	0.184	0.045*	0.032	0.012	-0.077
	(0.330)	(0.257)	(0.025)	(0.022)	(0.132)	(0.084)
hhhres	-0.008	-0.115	-0.006	-0.002	-0.084	-0.022
	(0.116)	(0.164)	(0.010)	(0.014)	(0.052)	(0.065)
lnhhadurbl	-0.257***	-0.035	-0.019**	-0.003	-0.004	-0.017
	(0.074)	(0.097)	(0.007)	(0.008)	(0.038)	(0.036)
y2013	0.588***	-0.173	0.004	-0.030	0.318***	0.372***
	(0.191)	(0.232)	(0.018)	(0.021)	(0.075)	(0.087)
y2015	0.111	-0.396	-0.015	-0.059***	0.531***	0.671***
	(0.218)	(0.247)	(0.018)	(0.021)	(0.085)	(0.101)
_cons	4.581	-1.651	1.014*	-0.395	-0.790	-6.220**
	(7.947)	(4.826)	(0.613)	(0.388)	(4.080)	(2.450)
Observation	2020	1986	2020	1986	2059	1962
Individuals	1197	1109	1197	1109	1161	1040
R^2	0.042	0.042	0.031	0.038	0.119	0.155
$\sigma_{\eta_{i1}}$	3.033	3.267	0.203	0.209	2.088	2.154
$\sigma_{v_{it1}}$	3.115	3.825	0.288	0.344	1.166	1.391
ρ	0.487	0.422	0.332	0.269	0.762	0.706

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for individual-level means of the time-varying exogenous variables.

(columns 1-9 in Table 2.13) and the probability of participating in none of these activities (column 10 in Table 2.13). We also estimate a bivariate random effects model for the retirement effect on the total number of activities that individuals take up (column 11 in Table 2.13). We estimate separately for men and women and report only coefficients of the key retirement variable to save space.

As shown in Table 2.13, transition into retirement significantly increases the probability of interacting with friends and going to a club (e.g., sports clubs or social clubs) for men, but reduces the probability of caring for a sick or disabled adult and the probability of doing none of the activities that we look at. The increases in social activities might explain their improvement in cognitive tests. The significant random effects in predicting both participation in social activities and transition into retirement suggest that bivariate random effects models should be estimated instead of pooled models. Covariances of random effects suggest that there are unobserved time-invariant individual effects that significantly predict the use of internet and the probability of staying in the labour force. Retirees are more likely to report less or zero number of activities. Probit random effects estimations show that male retirees overall are less likely to play games, participate in community activities, do voluntary work, and care for sick adults.

As for women, bivariate RE models' estimations suggest that there are insignificant or marginally significant changes in their probability of participating in social activities after retirement. The lack of compensating social or intellectual activities after the loss of work-related simulations might explain the short-run decline in their cognitive tests' scores after retirement. Female retirees overall are less likely to provide help to others, participate in community activities, invest in stock markets, and use Internet.

2.8 Conclusion

This paper examines both the short-run effect of entering formal retirement for urban formal workers reaching the mandatory retirement age, and the cumulative effect of years in retirement for these retirees on their cognitive functioning, depression risks, and physical well-being. We identify the effect of formal retirement using mandatory retirement ages for blue-collar and white-collar workers as instruments, and separate analyses for men and women. We use data from a nation-wide, longitudinal household survey and restrict the sample to formal workers or formal retirees who have urban residence permits (hukous) and live in urban areas. We use a bivariate random effects model to study changes in mental and physical health outcomes after transition into retirement or with years in retirement. The model accounts for the binary nature of the endogenous retirement status by using probit random effects model to estimate the retirement transition. It also accounts for the unobserved time-invariant individual effects that determine both health outcomes and retirement by allowing for

Table 2.13: Estimation of Retirement Effect on Participation in Social Activities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Men Bivariate RE											
retire	0.328**	0.283	-0.033	0.299*	-0.022	-0.447	-0.698*	0.102	0.308	-0.353**	0.147
	(0.145)	(0.200)	(0.177)	(0.155)	(0.233)	(0.283)	(0.382)	(0.405)	(0.225)	(0.170)	(0.113)
var(η_{i1})	0.699***	3.233***	0.384***	0.716***	0.837***	0.841***	0.580**	4.791***	2.872***	0.940***	0.757***
	(0.113)	(0.456)	(0.110)	(0.127)	(0.221)	(0.324)	(0.296)	(1.600)	(0.503)	(0.155)	(0.054)
var(η_{i2})	4.674***	4.675***	4.591***	4.635***	4.633***	4.680***	4.592***	4.685***	4.668***	4.581***	4.632***
	(1.049)	(1.062)	(1.024)	(1.040)	(1.030)	(1.065)	(1.036)	(1.063)	(1.029)	(1.024)	(1.031)
cov(η_{i1}, η_{i2})	-0.302	-0.242	-0.328	-0.338*	-0.449	-0.180	0.454	-0.408	-0.980***	0.450**	-0.430***
	(0.186)	(0.294)	(0.225)	(0.199)	(0.313)	(0.338)	(0.507)	(0.586)	(0.341)	(0.217)	(0.160)
Observation	3445	3445	3445	3445	3445	3445	3445	3445	3445	3445	3445
log likelihood	-2811	-2582	-2023	-2475	-1600	-1277	-1214	-1359	-1985	-2494	-6131
Men RE											
retire	0.157	0.176	-0.235**	0.107	-0.278*	-0.555***	-0.414**	-0.089	-0.144	-0.094	0.088
ln[var(η_{i1})]	-0.360**	1.175***	-0.992***	-0.340*	-0.203	-0.171	-0.564	1.369***	1.022***	-0.072	-0.076
	(0.163)	(0.139)	(0.298)	(0.184)	(0.262)	(0.323)	(0.594)	(0.299)	(0.178)	(0.167)	(0.167)
Observation	3138	3138	3138	3138	3138	3138	3138	3138	3138	3138	3138
Individuals	1775	1775	1775	1775	1775	1775	1775	1775	1775	1775	1775
log likelihood	-1956	-1726	-1168	-1620	-745	-421	-359	-502	-1134	-1640	-1640
Women Bivariate RE											
retire	0.075	0.349	-0.296*	0.216	-0.131	-0.383	0.057	-0.414	-0.403*	0.002	-0.144
	(0.140)	(0.245)	(0.159)	(0.179)	(0.239)	(0.319)	(0.293)	(0.361)	(0.244)	(0.178)	(0.127)
var(η_{i1})	0.467***	3.522***	0.403***	1.182***	1.010***	1.853***	0.766**	2.703***	2.468***	0.932***	0.943***
	(0.086)	(0.532)	(0.114)	(0.175)	(0.261)	(0.608)	(0.318)	(0.844)	(0.462)	(0.155)	(0.064)
var(η_{i2})	6.651***	6.656***	6.630***	6.636***	6.589***	6.652***	6.527***	6.600***	6.612***	6.659***	6.669***
	(1.759)	(1.753)	(1.756)	(1.756)	(1.738)	(1.758)	(1.742)	(1.746)	(1.733)	(1.761)	(1.761)
cov(η_{i1}, η_{i2})	-0.003	-0.306	0.055	0.059	-0.249	-0.064	-0.306	-0.289	-0.477	0.040	-0.126
	(0.188)	(0.388)	(0.205)	(0.259)	(0.324)	(0.411)	(0.387)	(0.507)	(0.381)	(0.248)	(0.188)
Observation	3086	3086	3086	3086	3086	3086	3086	3086	3086	3086	3086
log likelihood	-2709	-2293	-1979	-2502	-1513	-1212	-1240	-1219	-1834	-2380	-5855
Women RE											
retire	0.073	0.212	-0.265**	0.246**	-0.264*	-0.417	-0.120	-0.556**	-0.632***	0.024	-0.046
	(0.094)	(0.181)	(0.103)	(0.119)	(0.153)	(0.254)	(0.173)	(0.233)	(0.168)	(0.126)	(0.127)
ln[var(η_{i1})]	-0.761***	1.262***	-0.909***	0.167	0.006	0.605*	-0.302	0.966***	0.889***	-0.071	-0.074
	(0.182)	(0.152)	(0.299)	(0.148)	(0.247)	(0.337)	(0.486)	(0.325)	(0.188)	(0.168)	(0.168)
Observation	2901	2901	2901	2901	2901	2901	2901	2901	2901	2901	2901
Individuals	1607	1607	1607	1607	1607	1607	1607	1607	1607	1607	1607
log likelihood	-1891	-1475	-1162	-1684	-695	-394	-423	-402	-1017	-1562	-1558

Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors. We also control for age, age square, marital status, education level, household size, survey year dummies, individual-level means of the time-varying exogenous variables. (1) Interacted with friends (2) Played Ma-jong, played chess, played cards, or went to community club (3) Provided help to family, friends, or neighbors (4) Went to a sport, social, or other kind of club (5) Took part in a community-related organization (6) Done voluntary or charity work (7) Cared for a sick or disabled adult (8) Stock investment (9) Used the Internet(10) None of these (11) Total number of activities.

the random effects of the health equation and of the retirement equation to correlate with each other. We also estimate single random effects models that do not account for potential endogeneity of retirement variables and compare the cognitive functioning and mental or physical health of retirees and formal workers, and of retirees having stayed for different years in retirement.

The relatively young retirement age of 50 for the majority of female workers and the retirement age of 60 for white-collar male workers mean that with the growing life expectancy, formal retirees now spend more time in retirement and in home environments. Given the long duration of their retirement, how their cognitive functioning declines, and subjective well-being changes will affect the potential economic costs coming from public transfers, medical and health care expenditures, opportunity costs of families and friends taking informal care and the let-go economic benefits of a healthy, older population. Older people in good health conditions can make valuable contributions to society by means of consumption, or unpaid labour such as caring for grandchildren and volunteering.

We find that transitions into retirement are endogenous to men's scores in cognitive tests as men scoring lower in the tests are also more likely to retire. Retirement is also endogenous to physical health problems for both men and women, as people reporting more physical limitations are more likely to retire. After accounting for the endogeneity of retirement decisions, we find a substantial and positive effect of transition into formal retirement on men's scores in cognitive tests, but a negative effect on women's scores in the mental intactness test. The gender-specific difference in the retirement effect may be explained by the more social activities men take up after retirement compared to women do. There is no significant retirement effect on either men's or women's subjective or physical well-being, as measured by the the number of depressive symptoms and the number of physical problems. The positive short-run effects of retirement on cognitive functioning for men disappear with the number of years in retirement and the negative effect for women reverses with more years in retirement. The progressive effect of retirement on cognitive ageing might be explained by the younger retirement ages for women. Again, the risks of depression and physical frailty are not found to significantly change with the years individuals spend in formal retirement.

The findings of this paper suggest that when making decisions on raising retirement ages, policy makers should also pay attention to the process of transition into retirement. They can introduce incentives for retirees to adopt a healthy lifestyle, support cognitive maintenance activities, or remove barriers that prevent retirees from taking part-time or informal jobs. The policy of delaying retirement may only postpone the transition problems. Occupational training programmes targeted at older workers may also help them maintain cognitive abilities and feel meaningful and useful. Active ageing programmes such as the Active, Connected, Engaged (Beedie et al., 2014) in the UK provide a lost-cost, yet sustainable physical intervention that relies on peer education may significantly reduce cognitive decline, prevent depression, and promote a healthier and rewarding retirement.

Chapter 3

Depression, Physical Health, Labour Market Exits and Entries of Older Informal Workers

3.1 Introduction

China has experienced rapid demographic transition in the past few decades and is facing the growing challenge of population ageing, as a result of decreased fertility and an increase in life expectancy. The one-child policy introduced in 1979 has caused the overall fertility rate to decline from 6.30 per women in 1968 to 1.66 per women in 2015, when the policy was removed. The fertility rate remains low due to urbanization and increasing female labour participation. At the same time, life expectancy in China has been significantly prolonged and increased from 43.73 in 1960 to 76.91 in 2019. The proportion of population aged 65 years or above in China has changed from 4.4% in 1950 to 12.0% in 2020, and is expected to climb to 20.7% in 2035 and 26.1% in 2050 (United Nations, 2019). The rapid growth of the older population has posed challenges to the pension system, state health insurance and other social security programmes. The challenge also shows regional differences with less developed and rural areas facing bigger threat to the sustainability of their social pension systems. Although the government has started to raise formal retirement ages, it does not affect the retirement decision of informal workers, who are not affected by the retirement ages and have already worked into very old ages. How labour supply in the old age affects health care expenditures and to what extent health conditions affects labour supply decisions are important in understanding the impacts of raising retirement ages and evaluating various costs of population ageing.

This chapter studies the dynamic relationships between health, both mental and physical health, and labour supply of older, informal workers in low socio-economic status (SES), using three waves (2011-2015) of data from the China Health and Retirement Longitudinal Studies (CHARLS). Health is the major source of uncertainty facing older workers and affects their labour force transitions (Bound et al., 1999; Haan and Myck, 2009). The relationships between physical or mental health and voluntary retirement (Disney, Emmerson and Wakefield, 2006), or involuntary job loss (Stewart, 2001) have been studied extensively, focusing on formal workers in the context of developed countries (García-Gómez, Jones and Rice, 2010). There is less evidence on the impact of health on labour market entries and exits (García-Gómez, Jones and Rice, 2010), which is common among older, informal workers. In contrast, formal retirees' retirement is more an absorbing state and is not usually followed by re-entering the labour market.

There is little evidence on the relationship between health and labour supply for older, informal workers from developing countries or in low-welfare, low socio-economic status. These informal workers can work in agricultural sectors, either employed or doing household agricultural work, can engage in self employment, or work in non-agricultural sectors under temporary or flexible contracts and thus not entitled to the retirement pensions, health insurance or unemployment compensation that are tied with employers. While older formal workers in developed countries can choose from a wide range of optionssto exit the labour force such as early retirement and disability benefits, the older, informal workers from developing countries tend to rely on private savings and private transfers, mainly from adult children, for old-age support after exit from the labour market. The low labour income levels, imperfect credit markets, limited coverage of social security programmes and health insurance, the large size the informal sector in developing countries and also in less developed areas of China, have seen informal workers work, either full-time or part-time, until very old ages, or till they are physically incapable of carrying on. The limited coverage of the social safety net, the low-quantity and low-quality health facilities in their residential places mean that the older, informal workers are particularly vulnerable to health shocks, which have been found to be closely related to exit from the labour market and transition into disability (García-Gómez, 2011) for workers in the European countries. Instead of studying health shocks or the first onset of severe chronic diseases/ functional limitations, we focus on health declines, both physical and mental health, and the health effects on older, informal workers' labour supply adjustment. We expect the health effect to be different for informal workers in developing countries because institutional difference in social security and health insurance generosity, labour market structure (Trevisan and Zantomio, 2016) can affect individual labour supply adjustments to health shocks. Older informal workers may exit the labour force after a health shock, but can re-enter the labour force shortly afterwards, as we show in this paper.

One contribution of the paper is that we study two-way causality of both health effects on labour market entries and exits, and the effect of labour market transition on various aspects of health. Most of the existing studies look at the one-way causality and either model health as a determining factor of retirement or unemployment, or examine the effects of retirement or unemployment on health outcomes (Haan and Myck, 2009; Clark and Oswald, 1994; Farré, Fasani and Mueller, 2018; Böckerman and Ilmakunnas, 2009). The commonly-used health measures that include physical well-being and mortality, and self-reported health. Recent studies start to investigate into mental health outcomes including cognitive functioning and depressive symptoms.

The second contribution of the paper is that we add the third dimension of mental health into the analysis of dynamic relationship between health and labour supply of older workers. Accounting for the correlation between physical health and mental health can reduce the unexplained variation in models of health, and provide a more comprehensive measure of the quality of life. The existing economics studies of individual health and labour supply look at only one aspect of health and fail to account for the fact that mental and physical health are interrelated, which has been supported by evidence from biological studies (Katon, Lin and Kroenke, 2007) and empirical economics studies (Ohrnberger, Fichera and Sutton, 2017*b*). Studying the interrelation between physical health and mental health also has implications on the spill-over effects of health policies.

We focus on depression symptoms such as sadness, pessimism, loss of agency, in studying mental health[1]. Depression is related to disability in later years of life, as suggested by epidemiologic and clinical studies (Bruce, 2001), especially for individuals who are aged above 70 and experience high level of depressive symptoms (Barry et al., 2009). Although the relationship might show modest gender difference (Bruce, 2001), the risk of depression is not gender-neutral. For example, Baranov et al. (2020) find the depression risk to be about twice as high for women than for men. The reasons might be that women tend to have lower bargaining power over household spending and take more informal caring jobs compared to men do. This paper therefore also investigates the gender-specific differences in health conditions and labour supply status, as well as their interrelationships. We define depressive symptoms as an indicator based on the 10-item Centre for Epidemiologic Studies Depression Scale, or the CES-D scores.

Physical health is measured by the total number of chronic diseases, limitations in daily activities, and limitations with mobility. The measure can better capture the accumulated burden of physical frailty than any single one of its elements. We define the labour market risk in multiple ways, focusing on the transition between working

[1] Depression symptoms are the second leading causes of disease burden worldwide (Ferrari et al., 2013). Its cross-sectional prevalence is estimated to be 4-5% of the global population (Vos et al., 2012), and its life-time prevalence is expected to be 13% using a sample of 18 countries (Kessler and Bromet, 2013).

(either in agricultural or non-agricultural sector) and non-working, the transition between agricultural jobs and non-agricultural jobs if they have been working for at least the first two waves. Formal retirees and the formal workers reaching the retirement ages are excluded from the sample because they are subject to the compulsory retirement policy and have to retire from their formal jobs whether they are in good or deficit health status upon the policy ages.

The third contribution of the paper is to estimate a dynamic cross-effects model and study the dynamic relationships between depression, physical health and labour market transitions. Most of the existing studies remain cross-sectional and provide evidence on one particular time period. The relationship between mental health and physical health, or labour supply status can vary overtime due to adaptation[2]. The likely correlations between health or labour supply variables and the random effects would bias the estimates, therefore we estimate a trivariate, non-linear, random effects model where the individual random effects of the three endogenous variables of depression risk, physical health problems, and labour supply risk are assumed to follow a joint normal distribution. Regressing on the lag of other health or labour supply variables can relieve the concern of reverse causality. We report the level of correlations between the unobserved time-invariant individual heterogeneities of the three processes and explain what the unobservables could be.

We find that working in the non-agricultural sector significantly reduces the risk of depression in the subsequent period for men. Depression increases the probability of re-entering the labour force in the following period for men, most likely out of financial stress, but it does not affect labour supply behaviours of women. An increase in the number of physical health problems predicts a higher risk of depression in the subsequent period. The effects are larger for men than for women, even after accounting for the unobservables, such as medical complications or stress in patients having certain chronic diseases, or genetic factors that increase risks of both depressive syndromes and other chronic diseases, health behaviours, lifestyle choices and social capital. These factors have been studied before (Katon, Lin and Kroenke, 2007; Bruce, 2001; Taylor, Aizenstein and Alexopoulos, 2013; Ohrnberger, Fichera and Sutton, 2017*b*). Depression does not have any effect on men's reported number of physical problems in the subsequent period but increases women's reported number of physical problems.

Physical health problems predict an exit from the labour force in the subsequent period, and the effect is larger for men than for women. The gender-specific differences can be explained by the fact that men take more labour-intensive jobs and are affected

[2]Economics studies of the happiness-health relationship has highlighted the role of adaptation, and that people's expectations for health standards affect their self-reported health and happiness (Graham, 2008). Individual expectations for health standards can change after experiences of chronic diseases.

more by physical limitations compared to women, who may spend more time taking care of families. There is no evidence of labour supply status affecting physical health in the subsequent period for either men or women. Nonetheless, women doing agricultural work and especially non-agricultural work in the previous period report significantly less physical problems than women who were not working do. Physical health deterioration does not have any significant effect on the probability of transition between agricultural and non-agricultural jobs for either men or women. We also find strong and significant state dependence for both depression, physical health problems and labour supply status, with gender-specific variations in the magnitudes.

The paper is organized as follows. The next section summarizes the relevant studies and the conceptual framework for explaining the interrelationships between depression, physical health, and labour supply behaviour. Section 3 introduces the data and the sample, the dependent and the explanatory variables we use for the estimation, and presents the summary statistics. Section 4 presents the econometric models, followed by the empirical results (Section V5). Section 6 provides the robustness checks, and section 7 concludes and discusses the policy implications of our findings.

3.2 Relevant Literature

3.2.1 The Dynamics of Health and Employment

Employment is state-dependent because prior work experiences and human capital investment acquired on the job training can affect current work choices (Heckman, 1981*a*). The duration of unemployment or being economically inactive, and the fixed costs incurred by labour force entrants affect the chance of individuals returning to work, generating structural state dependence. Working may also alter individual preferences and reinforce labour supply behaviour (Heckman, 1981*a*). Higher search costs, stigma effects or human capital depreciation among the non-employed can explain state dependence of non-working (Hyslop, 1999).

Health is state-dependent due to the long-term nature of health conditions, stable demand for health and health care, or chronic nature of some diseases (Haan and Myck, 2009). There are some studies using longitudinal data and estimating the dynamics of health (Contoyannis, Jones and Rice, 2004; Hauck and Rice, 2004; Contoyannis and Li, 2011). For example, Hauck and Rice (2004) study the level of mental health mobility in the UK and find that the lower-income and less-educated groups suffer from both greater mental ill-health and greater persistence. Blundell et al. (2016) estimate a dynamic model of health and employment for older workers in England and the US, and find that persistent health shocks have much bigger effects on employment than transitory health shocks. The total health impact on employment is much larger than what OLS estimates imply. Similar to Hauck and Rice (2004), they find a greater persistence of physical health problems among the low educated and the elder age groups. Accounting for state dependence can also improve the explanatory power of models of health. It has been found to explain 30% of the total unexplained variation in self-assessed health in Contoyannis and Jones (2004)'s study of health dynamics in the UK and 18% the unexplained variation in mental health in Hauck and Rice (2004). Health inequalities may be over- or under-estimated, depending on the level of persistence or health mobility, if they are measured at a single point in time in cross-sectional studies (Hauck and Rice, 2004).

Socio-economic status may also explain the persistence in health outcomes. Community epidemiological studies have found that women, older people, and individuals who are separated or divorced are all faced with higher prevalence of depression (Kessler and Bromet, 2013). Existing studies using data of older people in China also find that age, low levels of education, retirement, physical health problems, poor social support and financial problems are correlated with higher levels of depressive symptoms(Lei et al., 2014*a*). Education, income levels and unobserved community influences such as THE prices and quality of health care services, public health infrastructures, inherent healthiness (e.g. water, sanitation and air quality) are found to explain the physical health inequality among older people in China (Lei et al., 2014*b*).

3.2.2 Depression and Physical Health

Mental health is an integral and essential component of health. Depression symptoms such as sadness, pessimism, loss of agency, are important aspects of mental health, and relate to deteriorating physical health. Evidence from epidemiologic and clinical studies suggests a reciprocal, potentially spiraling relationship between depression and disability in later years (Bruce, 2001), especially for those aged above 70 (Barry et al., 2009). The Global Burden of Disease (GBD) 2010 study finds that depression is the second leading causes of disease burden worldwide (Ferrari et al., 2013). Its cross-sectional prevalence is estimated to be 4-5% of the global population (Vos et al., 2012), and its life-time prevalence is expected to be 13% using a sample of 18 countries (Kessler and Bromet, 2013).

Although the relationship between depression and disability shows modest gender differences (Bruce, 2001), the risk of depression is about twice as high for women as it is for men (Baranov et al., 2020). The reason might be the lower bargaining power of women over household spending and other issues such as more informal caring jobs they take.

Evidence from biological studies has supported bidirectional effects between depression or anxiety and the severity of medical illness. Depression can affect physical health directly through increasing medical complications, such as increasing muscle tension (Katon, Lin and Kroenke, 2007)[3]. Depression can affect physical health indirectly through health behaviours such as diet, exercise, cessation of smoking and medication. By distorting beliefs, depression leads to lower aspirations and lower investments on health and cognitive preservation (Dalton, Ghosal and Mani, 2016).

Explanations for the effect of physical health on depression focus on two aspects. The vascular depression hypothesis suggests that cerebrovascular diseases may predispose, precipitate or perpetuate some geriatric depressive syndromes, and vascular diseases have been found to share common genetic determinants with depression in old age (Alexopoulos et al., 1997; Taylor, Aizenstein and Alexopoulos, 2013). Moreover, chronic diseases can affect the integrity of subcortical regions of the brain involved in mood regulation and increase the risk of developing depression symptoms (Taylor, Aizenstein and Alexopoulos, 2013). From the psychological perspective, physical problems create stressful conditions that increase the risk of depressive symptoms by affecting individual social and emotional functions, such as maintaining productivity, social relationship and independence (Luo, Chui and Li, 2019; Bruce, 2001). Empirical evidence from socio-economic studies finds that chronic physical health problems are key risk factors of depression symptoms among elder population (Huang et al., 2010). Empirical evidence focuses on how physical health and mental health affect each other through individual engagement with the environment, including physical, cognitive, and social activities. Adopting the mediation analysis within the health economic framework of health production and consumption, Ohrnberger, Fichera and Sutton (2017*b*) provide evidence on the mediating effects of lifestyle choices and social capital. Specifically, they find that mediating effects, mainly previous participation in physical activities, account for about 10% of the total effect of past physical (mental) health on current mental (physical) health.

Previous studies tend to be cross-sectional and study correlations. For the few economics studies looking at the dynamic cross-effects between physical health and mental health, they tend to find a reciprocal relationship between physical health and mental health, though the effect is asymmetric with physical health having a larger impact on the subsequent mental health (Luo, Chui and Li, 2019; Ohrnberger, Fichera and Sutton, 2017*a*).

[3]A review from Katon, Lin and Kroenke (2007) finds that patients with chronic diseases and comorbid depression and anxiety report higher numbers of medical symptoms than those with only chronic diseases do, after controlling for severity of the chronic diseases.

3.2.3 Depression and Labour Market Participation

Theoretically, Baranov et al. (2020) explain how depression can make people less likely to return to work after being economically inactive or work less than they otherwise would do by affecting their preferences, expectations, and causes financial impairment. Depression can cause time discounting, which can lead to impulsivity, inconsistent intertemporal choices, and to present-biased behaviors such as drinking, smoking, or suicides. A second distortion on preference is through disutility or increasing costs of effort. The higher cost of effort associated with stress and fatigue for the depressed individuals can increase the mental cost of simple tasks such as physical activity and social engagement that are beneficial for physical health and cognitive functioning, and ultimately leading to withdrawal from the labour market. Depression affects expectations and may induce a pessimistic view about individuals' ability, or about the returns to work, making them less likely to return to work after being economically inactive, or tend to work less than they otherwise would do. Last but not least, depression affects constraints by increasing sick days and reducing productivity in jobs (Baranov et al., 2020). From these perspectives, depression is related to decreasing labour supply. Empirical evidence using data of employees in developed countries also supports the view that depression is associated with reductions in labour force participation and employment (Chatterji, Alegria and Takeuchi, 2011; Frijters, Johnston and Shields, 2014), and generates great costs for those economies (Sobocki et al., 2006; Stewart et al., 2003).

Alternatively, if individuals become depressed because of financial stress and/or unemployment, it has been shown that they are more likely to search intensively for jobs and re-enter the labour market (García-Gómez, Jones and Rice, 2010). Therefore, the effect of depression on labour market participation is an empirical question.

3.2.4 Labour Supply Decision and Health

Non-working can have negative effects on physical and mental health outcomes. First of all, exit from the labour market has monetary costs that can hardly be compensated for by the low retirement pension or unemployment pensions for informal workers in China and in other developing countries. Involuntary terminations usually cause a loss of long-term earnings (Ruhm, 1991), and force the economically inactive to adjust to lower quality of nutrition, housing, and health care that may result in deteriorating health outcomes (Farré, Fasani and Mueller, 2018). The financial stress associated with joblessness could result in mental ill-health. Furthermore, exit from the labour market usually means the loss of social networks and physical activities related to the work environment, changes in health behaviours such as physical activities, smoking and drinking behaviours. Working, especially in non-labour-intensive jobs, may help preserve cognitive functioning and a healthy mental condition (Llena-Nozal, Lindeboom and Portrait, 2004). Nonetheless, mental strain in jobs also affects health. Highly physically demanding jobs such as agricultural may cause physical health to deteriorate faster than in less labour-intensive jobs. Therefore, the overall health effect of being economically inactive is ambiguous in both literature and in theory, and empirical evidence is necessary for inference based on a specific setting.

3.3 Data and Summary Statistics

3.3.1 The China Health and Retirement Study (CHARLS)

We use three waves of the data from the CHARLS covering the period from 2011 to 2018. The CHARLS is a biannual, nation-wide longitudinal survey that collects information on the older people aged 45 years or above and their spouses in the baseline survey. The 2011 baseline survey collects information of 17,708 respondents in 10,257 households residing in 150 counties in 28 provinces[4]. Subjective and objective data are collected on individual health and well-being, individual and household socio-economic conditions, employment and retirement, health care and insurance. These respondents are followed-up in 2013, 2015 and 2018, though some of them dropped out in the middle of the surveys and are replaced by new participants.

We choose the CHARLS because it contains the most detailed and comprehensive information on older people's health status and functioning. It also provides tests of cognitive functioning and depressive symptoms. The population of individuals aged 45 and above, or the middle-aged and the elderly, are ideal for studying the relationship between health and employment, as they are faced with higher risks of physical and mental health problems than the younger population.

The *Gateway to Global Aging Data* have published the harmonized CHARLS dataset and codebooks to facilitate cross-country comparisons. The paper uses data of the Version C of the Harmonized CHARLS dataset, and supplements with some variables from the original dataset but not appearing in the Harmonized dataset (e.g. community survey data, and detailed information about jobs' characteristics).

3.3.2 Outcome and Control Variables

This section introduces the health and labour supply outcomes variables, as well as the SES we control for in empirical models. Table 3.1 summaries the abbreviations of these variables and their definitions.

[4]Tibet is not included in the survey, and two other provinces of Hainan and Ningxia are also not represented in the study due to their relatively small sizes of population

3.3.2.1 Depressive Symptoms

Depressive symptoms are measured by the Centre for Epidemiologic Studies Depression Scale, or the CES-D scores. It is developed by Radloff (1977) and has become one of the most widely used tests to diagnose depression quotient. The standard CES-D measure contains 20 items and the CHARLS survey uses a simplified, 10-item CES-D measure. Specifically, respondents are asked about the frequency in which they experience a certain feeling over the week prior to the interview. The 10 items cover 8 negative feelings of feeling depressed, feeling that everything was an effort, restless sleep, feeling lonely, bothered by little things, feeling that they could not get going, feeling it hard to concentrate, feeling fearful; and 2 positive feelings of feeling happy and feeling hopeful. The 4 options of frequency are 'almost never (less than one day)', 'sometimes (1–2 days)', 'often (3–4 days)', and 'most of the time (5–7 days)', and are coded as 0, 1, 2, 3 for negative feelings and 3, 2, 1, 0 for positive feelings. The total CES-D score ranges from 0 to 30, with a lower score indicating a better mental health condition.

Previous studies have generally selected a cut-off score of 10 in the 10-item CES-D Scale which ranges from 0 to 30, and found it to have high specificity in older samples (Andresen et al., 1994; Chen and Mui, 2014). Lei et al. (2014*a*) using CHARLS to examine depressive symptoms and SES among the elderly in China also adopt the cut-off point of 10. We construct an indicator of depression that equals 1 if an individual scores 10 or above in the CES-D test. We use the categorical variable instead of the original CES-D scores because the scores measure subjective well-being which can be transitory and affected by the respondents' emotional conditions during the interview. Apart from the concern over measurement errors, we want to focus on old-age depression, which is more persistent and causes more severe problems than feeling bad or bored occasionally.

3.3.2.2 Physical Health Measures

Following the literature, we define physical frailty as an accumulated burden of chronic diseases, functional limitations, and other health-related deficits and symptoms. Specifically, we construct the physical health variable by including 12 kinds of chronic diseases, 6 kinds of limitations in activities of daily living, and 7 kinds of limitations with mobility[5][6][7]. Physical health measure is the sum of the number of chronic diseases

[5]Chronic diseases include hypertension, diabetes, cancer or a malignant tumor, chronic lung disease, heart problems, stroke, arthritis, dyslipidemia, liver disease, kidney disease, stomach or other digestive problems, asthma.

[6]The ADLs included are dressing, bathing and showering, eating, getting in and out of bed, using the toilet and controlling urination and defecation.

[7]Mobility tests include running or jogging for 1km, walking for 1km, walking for 100 meters, getting up from a chair after sitting for long periods, climbing several flights of stairs without resting,

or limitations, with a range between 0 and 23. A higher score indicates worse physical health.

The measure can better capture the accumulated burden of physical frailty than any single one of its elements as individuals with few mobility constraints or certain types of chronic diseases may still be able to carry on with their work. The distribution of the number of ADLs is heavily skewed to the left with more than 75% reporting no limitation[8].

3.3.2.3 Labour Force Status

We define three types of labour force status based on the self-reported types of employment. The harmonized survey published by Global Ageing Survey Data summaries an individual's main labour force status as agricultural work, non-agricultural employed, non-agricultural self-employed, non-agricultural unpaid family business, retired, and never worked[9][10]. If respondents engage in more than one types of jobs, they are categorized into the job type that they spend the most time on, or to non-agricultural work if they spend the same time on agricultural and non-agricultural jobs.

There is a substantial number of respondents who report being retired in earlier waves who then re-enter to the labour force in later waves, especially for those reporting to be retired from agricultural jobs. We regard it more accurate to define these people as being economically inactive or not working, instead of being retired. The major difference is that retirees receive retirement pensions and not usually go back to work after retirement, while the economically inactive may receive no or little pensions and are more likely to go back to work for example when they are faced with financial stress or recover from a health shock.

We create an indicator variable for the economically inactive that equals 1 for individuals self-reporting as being not currently working but not being unemployed, and 0 for individuals currently working, either in agricultural or non-agricultural jobs. To study state dependence, we create an indicator variable that equals 1 for individuals who take agricultural jobs and 0 for individuals taking non-agricultural jobs.

---

stooping kneeling, or crouching, lifting or carrying weights over 5kg, picking up a coin from the table, and reaching arms above shoulder level.

[8]There is a non-trivial amount of respondents reporting certain kind of ADLs in one wave but no limitation at all in the following wave. The lack of persistence that usually features severe physical problems makes the number of ADLs a noisy measure of physical health if being used alone.

[9]Agricultural work includes individuals doing agricultural work for their own families or others in wage for at least 10 days

[10]Respondents who have worked for at least three months during their lifetime and have searched for a new job during the last month are categorized as unemployed, otherwise as retired if they have not searched for a new job during the last month.

To study the lagged effects of the different labour force status on health, we create an indicator of agricultural work that equals 1 for agricultural workers, 0 for non-agricultural workers or the economically inactive. Similarly, an indicator of non-agricultural work is created which is equal to 1 for non-agricultural workers (including the non-agricultural employed, the non-agricultural self-employed, those doing non-agricultural unpaid family business), and equalling 0 for agricultural workers or the economically inactive. The reference group in the health equations would be individuals not working in the previous period.

3.3.2.4 Control Variables

To account for factors that affect both health and retirement decisions of individuals, we control for a large set of socio-economic status variables, time and location dummies. The demographic variables contain an indicator for rural hukou-holders, age and age squared, indicators for age groups, educational levels, and marital status. Hukou is the legal residence in China and rural hukou-holders tend to receive less benefits in terms of health care, social welfare and education than urban hukou-holders. Household-level variables include the number of children, the number of household members, and the logged values of household durable assets. We use the logged values of household durable assets to measure the household wealth because it is a better measure of long-run resources. Measures of household incomes and expenditures in the survey show greater variation, and also have more missing data.

We also control for residential places of the respondents using an indicator of whether living in urban areas and seven area dummies (area1-7) defined by the National Statistics Bureau. Among them, urban areas include city-centres, city-towns, small city-centres, small city-towns and special areas, and rural areas include towns and villages. Provincial difference can play an important role in the level of health inequality and the health–socio-economic-status gradient because of health care and other prices, inherent healthiness of the area, public health infrastructure and other factors (Lei et al., 2014*b*). Therefore, we control for provincial differences in health and labour market conditions by including 28 provincial dummies.

Finally, indicators of the survey year of 2013 and 2015 are also included to account for the time trends of aggregate health and labour supply behaviour, time-varying reporting changes and effects of age that are not captured by age and its quadratic term.

Table 3.1 shows short-forms of the dependent and explanatory variables in our empirical models and their definitions.

Table 3.1: Definitions of Covariates

rrural	1 if having a rural hukou, 0 otherwise
rage	Age defined by birth year on ID, and reported age if missing ID information
rage_sq	Age square/100
rage5055	1 if 50 <= $rage$ < 55, 0 otherwise
rage5560	1 if 55 <= $rage$ < 60, 0 otherwise
rabove60	1 if $rage$ >= 60, 0 otherwise
rnprimary	1 if no primary school degree, including illiterate, not finishing primary school, only receiving private education (sishu), 0 otherwise (taken as reference group in the models)
rprimary	1 if highest educational attainment is finishing primary school, 0 otherwise
rsecondary	1 if highest educational attainment is finishing secondary school, 0 otherwise
rhighabove	1 if highest educational attainment is finishing high school or above (including vocational school, colleges, universities (bachelor, master or phd degree), 0 otherwise
rmarried	1 if married or partnered, 0 otherwise
hchild	number of living children
hhhres	number of people living in this household
lnhhadurbl	log (household durable assets' values)
y2013	1 if Wave 2 (2013)
y2015	1 if Wave 3 (2015)
urban_nbs	1 if residing in urban areas (including area 1-5), 0 if residing in rural areas (including area 7 and 8, categorized by National Bureau of Statistics)
area1	1 if residing in city-centre (area id= 111 in National Bureau of Statistics), 0 otherwise (taken as reference group in the models)
area2	1 if residing in city-town (area id= 112), 0 otherwise
area3	1 if residing in small city-centre (area id= 121), 0 otherwise
area4	1 if residing in small city-town (area id= 122), 0 otherwise
area5	1 if residing in special area (area id= 123), 0 otherwise
area6	1 if residing in town centre (area id= 210), 0 otherwise
area7	1 if residing in village (area id= 220), 0 otherwise

3.3.3 Sample and Summary Statistics

We restrict the sample to individuals who were aged 40 or above in the first wave, stay for at least wave 1 and 2, and also meet the below criteria. Firstly, they should not miss data on the health and labour supply outcome variables. Secondly, they have not ever reported to have completed internal or formal retirement process throughout the survey periods. They have not ever been categorized as formal workers reaching normal retirement ages during the survey periods[11]. Because we cannot identify whether individuals are formal or temporary workers, we use a rough screening method and define formal workers by the types of employers they reported that they work for. We define people working for government, public institutions, non-governmental organizations (NGOs) and firms as formal workers, and assume they are all subject to the compulsory retirement ages. This is a strict restriction as individuals working in these firms or institutions might also be informal or temporary workers not signing a long-term contract and are thus not subject to the compulsory retirement policy.

The reason that we exclude from our sample the formal retirees and the formal workers reaching the retirement ages is because they are subject to compulsory retirement policy, whether they are in good or deficit health status upon retirement. In other words, retirement is not a choice for them.

Table 3.2 gives a summary statistics of outcome and control variables based on our estimating sample. Women report worse mental (M_{it}) and physical health (P_{it}) status than men do across all labour force status, although the average ages are similar for both gender groups. The proportion of agricultural workers (L_{it}^{F}) are similar for men and women, while the proportion of non-agricultural workers (L_{it}^{NF}) doubles for men compared to it is for women, meaning men are more likely to work in the non-agricultural sector than women in our sample do. Women are more likely to be not working (L_{it}^{O}) than men do across all waves, indicating the different family roles they take compared to men. The retirement rate remains below 30% for men across waves. Note that because formal retirees are excluded from the sample, the economically inactive group mainly comprises individuals who previously worked in the informal sectors. The informal workers tend to have lower pensions and less generous health insurance coverage than the formal workers, and are more likely to work until they are physically incapable of doing so.

Educational attainment might have large, positive health effects. Women overall receive less education than men do, which is consistent with the gender-specific inequality of health. Ageing is also closely related to living without a partner, which may also contribute to the higher depression rate over time. Women are more likely to live without a partner, either because women tend to outlive their male partners, or because they tend to marry someone older than them and are thus more likely to become widow when they grow old.

Rural hukou serves as a proxy for individual birthplace or childhood environment. Rural-hukou holders might migrate to the urban areas and apply for urban hukous

[11] According to the official document issued by the State Council in 1978, normal retirement ages are 60 for white-collar male workers, 55 for blue-collar male workers, 55 for female civil servants or managers, and 50 for other female workers. Because we cannot identify whether they are white-collar or blue collar workers, we use the earliest possible retirement ages, which is 55 for men and 50 for women.

Table 3.2: Summary Statistics by Wave 2 labour Force Status and Gender Group

	Men			Women		
	farm	non-farm	not work	farm	non-farm	not work
depress	0.32	0.19	0.39	0.46	0.35	0.46
phyhealth	2.91	2.14	4.84	3.60	2.84	5.27
	(2.81)	(2.36)	(3.91)	(3.01)	(2.70)	(3.88)
farm work(2015)	0.77	0.22	0.19	0.77	0.30	0.21
non-farm(2015)	0.11	0.67	0.07	0.05	0.48	0.04
not work(2015)	0.12	0.11	0.74	0.19	0.22	0.75
rural hukou	0.97	0.79	0.92	0.97	0.77	0.88
age	60.65	53.28	67.57	58.62	53.70	64.87
	(8.50)	(7.05)	(9.82)	(8.28)	(8.13)	(10.20)
no primary edu	0.23	0.11	0.26	0.20	0.19	0.20
primary edu	0.31	0.23	0.28	0.17	0.21	0.16
secondary edu	0.23	0.40	0.17	0.11	0.25	0.09
high edu or above	0.07	0.22	0.05	0.02	0.11	0.03
married	0.91	0.96	0.80	0.89	0.90	0.73
num. of child	2.75	2.20	3.42	2.80	2.42	3.39
	(1.35)	(1.11)	(1.72)	(1.28)	(1.29)	(1.63)
household members	3.48	3.63	3.39	3.60	3.63	3.34
	(1.70)	(1.55)	(1.97)	(1.75)	(1.65)	(1.82)
log(hh durables)	6.73	7.66	6.08	6.83	7.48	6.38
	(1.95)	(1.73)	(2.61)	(1.95)	(2.02)	(2.53)
urban areas	0.16	0.44	0.33	0.18	0.52	0.38
city centre	0.01	0.10	0.06	0.01	0.14	0.09
city-town	0.01	0.04	0.05	0.02	0.03	0.05
small city centre	0.06	0.19	0.13	0.07	0.24	0.15
small city-town	0.08	0.10	0.09	0.09	0.10	0.09
special area	0.00	0.01	0.01	0.00	0.01	0.01
town	0.05	0.05	0.04	0.04	0.04	0.05
village	0.79	0.51	0.63	0.79	0.44	0.57
N	6,476	2,142	2,174	8,781	1,167	4,513
Total		10,792			14,461	

Notes: Standard deviations for discrete variables are reported below the means. We select on individuals who aged 40 or above at the first wave, not missing data on the health and labour supply outcome variables for wave 1 and 2, not belonging to the group of formal retirees or formal workers around the retirement ages.

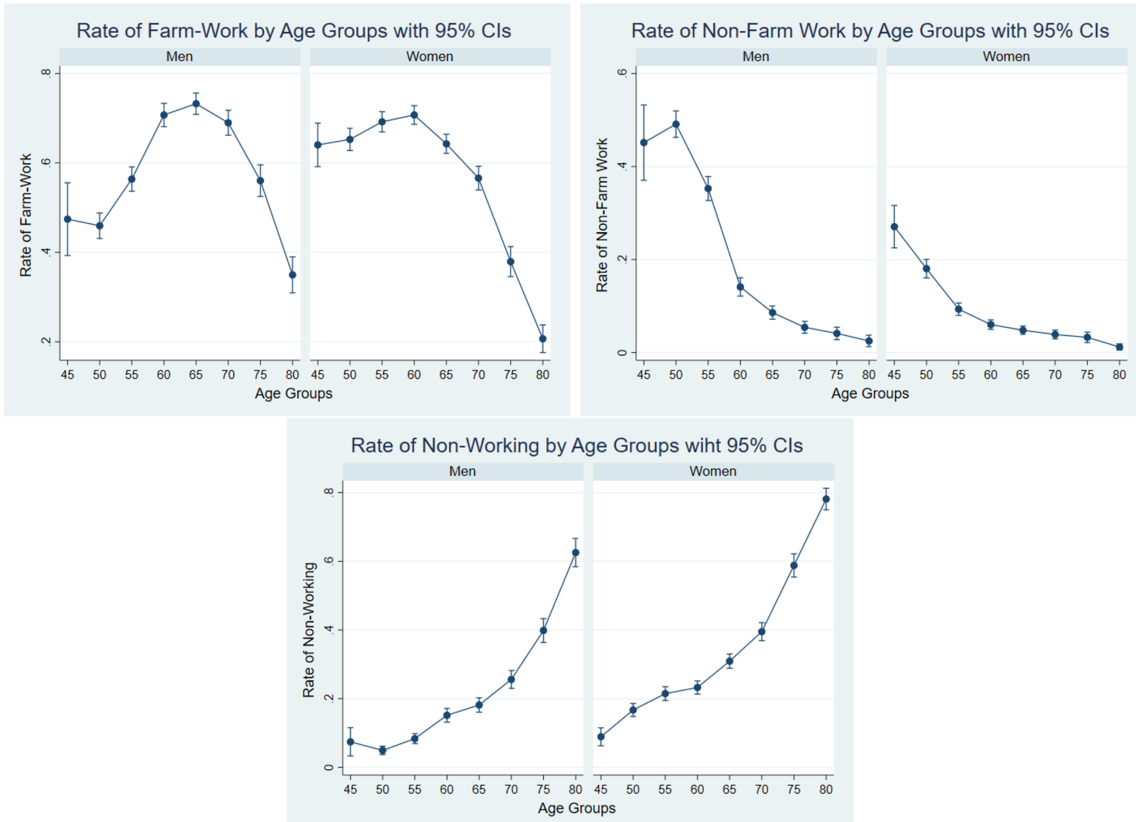

Figure 3.1: Labour Force Status by Age Groups

when they work or get married, but urban-hukou holders cannot apply for a rural hukou. Therefore, rural-hukou holders must be born in rural areas. The majority of our sample are rural-hukou holders. Holding a rural hukou means getting allocated lands (lands for housing, farming and forestry) from the state, and a high probability of doing agricultural work. We also control for individuals current living locations using the indicators of whether they live in urban areas and types of the areas they live in. There are some rural-hukou holders staying in the urban areas. They can be migrant workers living in the cities, more likely doing non-agricultural work, and still keep their rural hukous. It might also be the case that urbanization process sees their residential places being categorized as urban areas.

3.3.3.1 Distribution of Health and labour Force Status

To implicate on a life-cycle trend of depression risk, physical health declines and labour supply, we present distributions of health outcomes and labour force status by age and gender groups in Figure 3.1-3.4.

Figure 3.1 presents the proportion of agricultural workers, non-agricultural workers and non-workers for each age groups (5-year-gap between each age group), separately for men and women. Panel 1 (upper left) shows that agricultural workers tend to work into older ages, and the working rate starts to drop substantially only after 70, when they are probably physically incapable of working. Among them, men tend to stay in the labour force till older ages than women do. Panel 2 (upper right) shows that non-agricultural workers leave their jobs much earlier at the age between 50 and 60. Note that the reference group consists of both agricultural workers and the economically inactive, so they might not entirely stop working after leaving the non-agricultural jobs but take up agricultural jobs. This can be supported by the distribution of not-workers by age groups (the bottom panel). Rate of not-working remains low (around 20%) for men below 70 and women below 60, consistent with the pattern of the rate of agricultural work.

Life expectancy has increased in recent decades as a result of the improvements in medical care. Understanding the health consequences of the increase in longevity, especially if it causes an increase in morbidity, has important policy implications. Morbidity in older ages can affect both health care and social security expenditures, and labour supply of the older workers.

Figure 3.2 and 3.3 show the health trends by age, and the inequality of such effects across labour force status and gender groups. Depression rate are overall higher for women than for men, but does not show a clear trend with ages (Figure 3.2). Non-agricultural workers are less likely to be depressed than agricultural workers or non-workers throughout the life cycle, while the later two groups show similarly high risk of depression.

The confidence intervals (CIs) in the figure tell us about the variability of the reported health status. A high variability can be either due to small sample size (e.g. non-agricultural workers who work after 65, or non-workers aged below 50) or unobserved heterogeneity, or prevalence of negative mental health shocks.

The lower variability of the depression rate among the agricultural workers compared to the non-workers and non-agricultural workers means that agricultural workers are faced with similarly high risks of depression, regardless of the age groups they fall into. Risk of negative mental health shocks is higher for the non-workers.

The negative age effect on health is more pronounced in terms of the total number of physical health problems, including both chronic diseases and functional limitations, as shown in Figure 3.3. Agricultural workers do not report substantially worse physical health than non-agricultural workers, although non-workers report significantly more physical problems than the current workers. The pattern applies to both men and women. Women report more physical health problems than men do across labour force status and age groups. This is consistent with women's overall worse mental

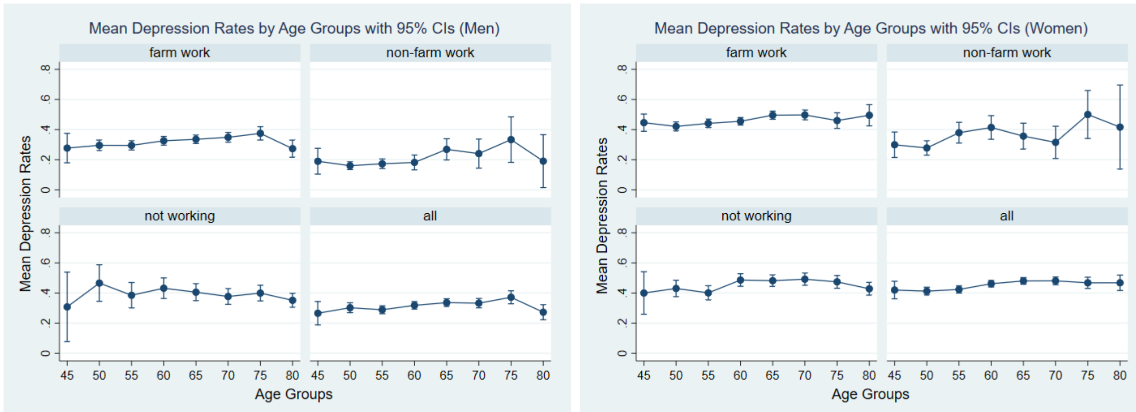

Figure 3.2: Depression Rate over Age Groups and labour Force Status

Figure 3.3: Physical Health Problems over Age Groups and labour Force Status

health conditions compared to men.

The variability of physical health, as shown by the sizes of confidence intervals, is larger among the non-workers than the others, suggesting that the elder non-workers are subject to more negative, physical health shocks apart from more physical health problems than others. High variability of physical health also implies a higher level of health inequality in the sense that some non-workers may remain healthy while others are faced with increasing health problems. The inequality is higher among male non-workers than female non-workers.

In the Appendix, Figure A.3.1-A.3.2 present the expectations and standard deviations of mental and physical health conditions by educational groups for men and women in different labour force status. In contrast to the age effects, individuals with higher educational attainment report better physical and mental health conditions, and are subject to lower variation. This applies to both gender across different labour

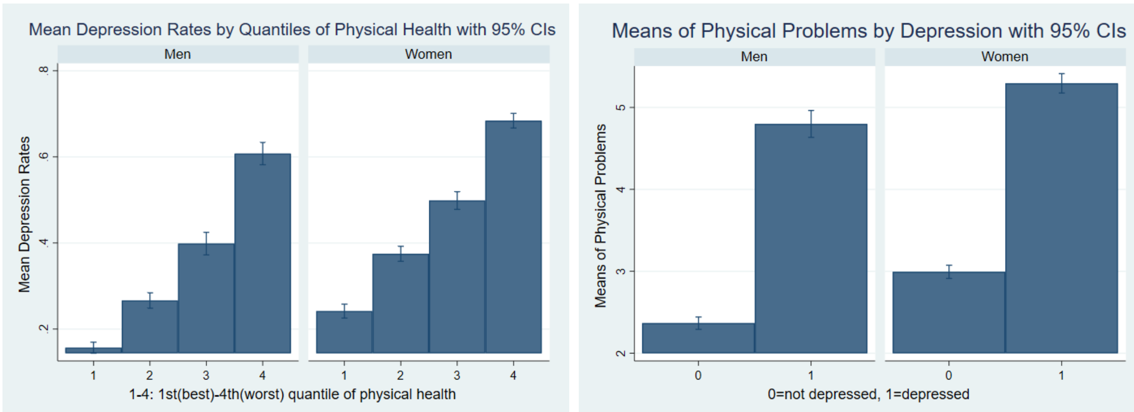

Figure 3.4: Correlations between Depression and Physical Problems

force status. Low-educated individuals are subject to worse mental and physical health conditions and higher risk of averse health shocks.

In summary, we find that physical health status declines with age and improves with educational levels, while depression risk only declines with education levels and not significantly varies by ages. Non-agricultural workers aged below 65 show the best mental and physical health conditions compared to others, and men better than women. Agricultural workers show a similarly high risk of depression but are in much better physical health conditions compared to non-workers, suggesting that factors other than physical problems may cause their depression symptoms, such as SES.

Depression rate increases with the number of physical health problems, as shown in Figure 3.4. We draw the expectations of depression rates across the four quantiles of physical health problems, separately for men and women. Women are more depressed than men are given the same level of physical problems, and report more physical problems than men do given that they are both depressed or non-depressed. The depressed individuals on average report two more physical problems than the non-depressed, similarly for men and women.

3.3.3.2 State Dependence

To shed a light on the persistence of health and employment outcomes, Table 3.3 to 3.5 report respectively the transition matrices of depression, physical health quantiles and labour force status between wave 1 and wave 2 or between wave 2 and wave 3. For example, the elements of the first row in the upper panel of Table 3.3 provide information about the distribution of depression in wave 2 given its distribution in wave 1. Note that persistence of certain state can either due to state dependence, individual socioeconomic characteristics, and the transition tables do not distinguish between them. We will identify the effect of state dependence in the empirical models in the following sections.

We use the balanced sample of individuals with no missing data on all three outcome variables for all three waves, and separately report the data for men and women. To simplify the table, we create four quantiles of physical health based on the number of physical health problems individuals report, so the lower quantiles correspond to the less physical problems and the better physical health status.

Table 3.3: Persistence of Depression for Across Waves

	Men			Women		
wave 1 \ wave 2	non-depressed	depressed	N	non-depressed	depressed	N
non-depressed	**82.64**	17.36	2,229	**74.55**	25.45	2,448
depressed	52.24	**47.76**	1,114	40.36	**59.64**	2,158
Total	58.53	41.47	3,343	68.92	31.08	4,606
wave 2 \ wave 3	non-depressed	depressed	N	non-depressed	depressed	N
non-depressed	**78.59**	21.41	2,424	**69.84**	30.16	2,696
depressed	43.42	**56.58**	919	30.99	**69.01**	1,910
Total	68.92	31.08	3,343	53.73	46.27	4,606

Notes: Percentages are reported in the table except for columns of N, which report the sample sizes. We select on the sample of individuals aged 40 or above at the first wave, not missing data on the health and employment outcome variables for all three waves, not belonging to the group of formal workers around compulsory retirement ages.

We find evidence of persistence in the two health outcomes and labour force status. Depression shows smaller persistence than physical health problems do. As shown in Table 3.3, only about 47% of men and 59% of women who were depressed in wave 1 remained depressed in wave 2, and only 56% of men and 69% of women who were depressed in wave 2 remained so in wave 3. Women tend to have higher persistence of depression, besides being more likely to be depressed compared to men.

Table 3.4 shows that people in the best and worst quartiles of physical health are more persistent than those in the 2nd or 3rd quartiles. And people in the 2nd or 3rd quartiles have a higher risk of moving towards the worse rather than the better physical health quartiles over time. Women suffer from worse physical health and higher risk of health declines as they are more persistent in the worst (4th) quantile, less stable in the best (1st) quantile, and more likely to transit into worse quantiles of health if they are in the middle of the health distribution. The risk of staying in the same quartile (besides the 1st quartile) or moving into the worse quartile of physical health is higher than the risk of being depressed, meaning that physical health might be more persistent than depression, though this again does not mean state dependence.

Table 3.4: Persistence of Physical Health Quantiles Across Waves

	Men					Women				
wave 1 \ wave 2	1st	2nd	3rd	4th	*N*	1st	2nd	3rd	4th	*N*
1st	**73.63**	19.19	4.95	2.24	1,475	**61.83**	25.96	9.28	2.93	1,433
2nd	18.98	**50.53**	21.75	8.74	938	18.25	**44.24**	23.12	14.39	1,293
3rd	6.71	28.57	**37.66**	27.06	462	5.68	26.73	**33.52**	34.08	898
4th	3.63	9.62	22.86	**63.89**	468	1.73	11.41	19.65	67.21	982
Total	39.25	27.94	16.69	16.12	3,343	25.84	28.14	20.10	**25.92**	4,606
wave 2 \ wave 3	1st	2nd	3rd	4th	*N*	1st	2nd	3rd	4th	*N*
1st	**64.33**	26.91	6.63	2.13	1,312	**55.21**	31.68	9.33	3.78	1,190
2nd	11.03	**55.78**	20.77	12.42	934	12.11	**38.50**	29.71	19.68	1,296
3rd	3.05	22.94	**36.02**	37.99	558	3.13	20.52	**34.34**	42.01	926
4th	0.37	4.45	18.37	**76.81**	539	0.59	5.19	13.48	**80.74**	1,194
Total	28.90	30.69	17.38	23.03	3,343	18.45	24.49	21.17	35.89	4,606

Notes: Ibid.

Table 3.5: Persistence of labour Force Status Across Waves

	Men				Women			
wave 1 \ wave 2	F	NF	NW	*N*	F	NF	NW	*N*
F	**80.66**	9.24	10.11	2,187	**81.96**	3.94	14.09	2,916
NF	25.24	**66.35**	8.41	630	30.03	**51.52**	18.46	363
NW	22.90	9.66	**67.44**	476	22.92	4.72	**72.35**	1,143
Total	61.71	20.22	18.07	3,293	62.44	8.05	29.51	4,422
wave 2 \ wave 3	F	NF	NW	*N*	F	NF	NW	*N*
F	**77.59**	10.96	11.45	2,044	**76.64**	4.62	18.73	2,813
NF	22.60	**66.91**	10.49	677	29.89	**47.01**	23.10	368
NW	19.15	7.53	**73.32**	611	21.84	4.11	**74.05**	1,337
Total	55.70	21.70	22.60	3,332	56.62	7.92	35.46	4,518

Notes: F stands for farm working, NF stands for non-farm working, NW stands for not working. Others are the same as above

In terms of labour force status (Table 3.5), agricultural workers in our sample show the highest persistence in staying in the same labour force status, followed by the economically inactive, and non-agricultural workers show the lowest persistence, similarly for men and women. Women are more likely to leave the labour force directly while men show higher chance of changing between job sectors.

3.3.3.3 Sample Attrition

We use an unbalanced sample of individuals staying for the first two waves of the survey, but not necessarily for wave 3. Health-related attrition due to deaths, serious illness or individuals moving into institutional care can lead to misleading estimation. The CHARLS data does not provide information on the reasons for leaving the survey except for whether the non-respondents are dead or still alive. Less than 30% of them are known to be dead (for whom a verbal autopsy is published in the 2013 survey), while the rest are either alive or dead. Individuals remaining in all sample periods may be systemically healthier and have lower rate of health decline than the population average (Contoyannis, Jones and Rice, 2004).

We compare summary statistics for the balanced and unbalanced sample to detect any significant attrition effects on the means of the variables. Table 3.6 shows that differences in means are very small, and we lose about 900 individuals in our balanced sample. There does not seem to be a significant health-related attrition. The unbalanced sample reports a similar risk of depression and worse physical health compared with the balanced sample.

3.4 Econometric Model

In this section, we present approaches to modeling the dynamics of depression, physical health problems, and labour force status. Before we study the sequential causality between health and employment, we need to account for the fact that both health and employment are persistent. Persistence in health and employment outcomes can be attributable to pure state dependence, observed individual socioeconomic characteristics, and unobservable, time-invariant individual heterogeneity. The dynamic random effect model enables us to distinguish between them and estimate the true level of state dependence for depression, physical health problems and labour supply status. Strong state dependence implies that health interventions and labour market policies can have long-run consequences and persistent effects. We control for the initial conditions using Wooldridge (2005)'s estimator and allow for potential correlations between the random effects and the time-varying exogenous regressors by adopting Mundlak (1978) and Chamberlain (1982)'s approach.

To study the sequential causality between depression, physical health and labour force status, we estimate a dynamic cross-effect model where the one-period lag of other health outcomes and labour force status are included as explanatory variables for health and labour market transition in the current period. Although reverse causality is not a concern in the cross-lag model that studies sequential causality, the major identification challenge comes from the many omitted variables that determine both health and labour market status, such as childhood health and other omitted socio-

Table 3.6: Summary Statistics by Waves Using Unbalanced and Balanced Samples

	wave 1		wave 2		wave 3	
	unbalanced	balanced	unbalanced	balanced	unbalanced	balanced
depress	0.42	0.41	0.36	0.36	0.40	0.40
phyhealth	3.13	3.09	3.54	3.46	4.40	4.32
	(2.97)	(2.93)	(3.18)	(3.09)	(3.65)	(3.57)
farm work	0.64	0.66	0.60	0.62	0.56	0.57
non-farm	0.13	0.13	0.13	0.13	0.14	0.14
not work	0.23	0.21	0.27	0.25	0.30	0.30
female	0.57	0.57	0.57	0.57	0.57	0.57
rural hukou	0.93	0.93	0.92	0.93	0.93	0.93
age	58.48	58.13	60.47	60.12	62.24	62.13
	(9.51)	(9.19)	(9.51)	(9.19)	(9.37)	(9.19)
no primary edu	0.21	0.21	0.21	0.21	0.21	0.21
primary edu	0.22	0.22	0.22	0.22	0.22	0.22
secondary edu	0.17	0.18	0.17	0.18	0.17	0.18
high edu or above	0.06	0.06	0.06	0.06	0.06	0.06
married	0.88	0.89	0.86	0.87	0.84	0.85
num. of child	2.79	2.75	2.88	2.85	2.98	2.97
	(1.42)	(1.39)	(1.45)	(1.42)	(1.45)	(1.44)
hh members	3.72	3.74	3.72	3.74	3.05	3.06
	(1.85)	(1.85)	(1.87)	(1.88)	(1.38)	(1.39)
log(hh durables)	6.73	6.77	6.80	6.85	6.75	6.75
	(1.99)	(1.95)	(2.22)	(2.18)	(2.23)	(2.22)
urban areas	0.26	0.25	0.26	0.25	0.26	0.25
city centre	0.04	0.04	0.04	0.04	0.04	0.04
city-town	0.03	0.03	0.03	0.03	0.03	0.03
small city centre	0.10	0.10	0.10	0.10	0.10	0.10
small city-town	0.09	0.09	0.09	0.09	0.09	0.09
special area	0.00	0.00	0.00	0.00	0.00	0.00
town	0.04	0.04	0.04	0.04	0.04	0.04
village	0.70	0.70	0.70	0.70	0.70	0.70
N	8,618	7,713	8,618	7,713	7,972	7,713

Notes: The unbalanced sample select on individuals aged 40 or above at the first wave, not missing data on the health and employment outcome variables for wave 1 and 2, not belonging to the group of formal workers around compulsory retirement ages. The balanced sample only differs from the unbalanced sample in selecting on individuals staying for all three waves.

economic status, preferences for health and working, community factors like prices and quality of health care and health inputs, environmental factors, local labour market conditions and economic development. These omitted variables would be correlated with the lags of health or employment, and other socio-economic status variables we control for, and bias the estimation. The likely correlations between health and labour supply effects and the random effects would make the cross-effects endogenous. We control for the initial condition and state-dependence of both the health or labour supply outcomes that we study by including them in the regressors, and control for the initial conditions and state-dependence of the other endogenous variables by estimating the trivariate dynamic random effects model and allowing for correlations between random effects of different outcomes.

All the empirical analysis is conducted separately for men and women to study gender difference. As the survey we use collects information of both respondents and their spouses (if present), by studying gender-specific pattern of health and labour market transitions, we also shed light on different family roles that men and women take. The assumption of zero correlation between random effects for different individuals will be violated due to spousal effects if we pool men and women in the regression. Due to the complexity of computing average partial effects for the trivariate, dynamic non-linear models, we focus on explaining the directions of the cross-effects between depression, physical health, and labour supply, and comparing the sizes of effects across gender and sectoral groups. We also study true and spurious state dependence, and compare the degree of true state dependence for different health and labour supply outcomes.

We start with the dynamic, nonlinear random effects model where cross-lagged effects of other dependent variables are not accounted for, and we focus on the persistence of health and labour force status. Then, we estimate the same dynamic random-effects model but include the lags of other dependent variables, to study the cross-lagged effects of health on labour market transitions, and how mental and physical health affect each other. Finally, we introduce the trivariate dynamic random effects model that the paper focuses on.

We use the dynamic random effects model rather than the Arellano and Bond (1991) model or other differencing methods to model state dependence, because the latter does not work for models with multiplicative unobserved effects, including binary choice models and count data models that our paper focuses on (Wooldridge, 2005). Also they tend to suffer from incidental parameters bias when the time periods are small, which is the case in our study.

3.4.1 Dynamic Non-linear Random-Effects Models

One advantage of the dynamic random effect model is the ability to distinguish between individual unobserved heterogeneity (random effects) and state dependence (the lagged dependent variable), both of which can explain the inter-temporal persistence of health outcomes and labour supply status. For binary outcomes of depression, and two indicators of labour force status (agricultural versus non-agricultural work, non-working versus working), we specify a dynamic, random-effects probit model. Specifically, let M_{it} be the measure of mental health, and it $= 1$ if individual i shows depressive symptoms at time t and $= 0$ otherwise.

Let $L_{it}^{W} = 1$ if individual i is doing agricultural work at year t, and $L_{it}^{W} = 0$ if doing non-agricultural work, and be missing for the non-workers. $L_{it}^{O} = 1$ if individual i is not working at year t, and $L_{it}^{O} = 0$ if if individual i is working (either doing agricultural work or non-agricultural work). Recall that we have excluded from the sample the unemployed, which is the minority in older people in our survey, and people who have never ever worked in their lives. The probability of depression can be specified as:

$$P(M_{it} = 1|M_{it-1}, ..., M_{i1}, \mathbf{x_i}, e_{1i}) = \Phi(x_{it}\beta_1 + \gamma_1 M_{it-1} + e_{1i}) \\ (i = 1, ..., N; t = 2, ..., T) \tag{3.1}$$

Similarly, the expectation of the probability of agricultural work versus non-agricultural work L_{it}^{W}, or the probability of non-working versus working L_{it}^{O}, can be specified as:

$$P(L_{it}^{W/O} = 1|L_{it-1}^{W/O}, ..., L_{i1}^{W/O}, \mathbf{x_i}, e_{3i}) = \Phi(x_{it}\beta_3 + \gamma_3 L_{it-1}^{W/O} + e_{3i}) \tag{3.2}$$

where x_{it} denotes the socio-economic variables, $\mathbf{x_i}$ is the row vector of x_{it} in all time periods $(x_{i1}, ..., x_{iT})$, M_{it-1} and $L_{it-1}^{W/O}$ are the one-period lags of the dependent variables. We assume that the dynamics are first order and follow a first-ordered Markov process as commonly assumed in literature. The individual effect e_i is additive inside the standard normal cumulative distribution function, and x_{it} are strictly exogenous. A lagged coefficient close to unity implies complete state dependence whereas a coefficient close to zero implies low state dependence. Lags or leads of x_{it} could have been included in the model, but we neglect the lagged effects of other socio-economic variables and focus on the interrelation of health and labour force status. Latent variable version of the above models can be derived as:

$$M_{it}^{*} = x_{it}\beta_1 + \gamma_1 M_{it-1} + e_{1i} + u_{1it} \tag{3.3}$$

where $u_{1it}|(M_{it-1}, ..., M_{i1}, \mathbf{x_i}, e_{1i}) \sim Normal(0, 1)$.

For a dependent variable that is a count variable, like the number of physical health problems, we specify a dynamic, random-effects Poisson model. Let P_{it}

be the number of physical health problems. It sums up the number of chronic diseases, limitations with daily activities (ADL), and limitations with mobility. Given $(P_{it-1}, ..., P_{i1}, \mathbf{x_i}, e_{2i})$, P_{it} follows a Poisson distribution with the mean given by:

$$E(P_{it}|P_{it-1}, ..., P_{i1}, \mathbf{x_i}, e_{2i}) = exp(x_{it}\beta_2 + \gamma_2 P_{it-1} + e_{2i}) \tag{3.4}$$

Nonetheless, in the above standard random-effects model, the individual-specific random effects (e_{1i}, e_{2i}, e_{3i}) are assumed to be uncorrelated with x_{it} in all equations.

3.4.2 Correlated Effects and Initial Conditions

To allow for the potential correlation between the random effects and the exogenous regressors, we follow Mundlak (1978) and Chamberlain (1982) and assume the relationship of the following form:

$$e_i = a_0 + \bar{x}_i a_1 + \alpha_i \tag{3.5}$$

where $\bar{x}_i$ is the individual-specific average value of x_{it} over the sample period. $\alpha_i \sim iidN(0, \sigma_\alpha^2)$ is independent of x_{it} and u_{it} for all i, t in all models. Thus the dynamic models above can be written as

$$P(M_{it} = 1|M_{it-1}, ..., M_{i1}, \mathbf{x_i}, \alpha_{1i}) = \Phi(x_{it}\beta_1 + \gamma_1 M_{it-1} + a_{10} + \bar{x}_i a_{11} + \alpha_{1i}) \tag{3.6}$$

$$E(P_{it}|P_{it-1}, ..., P_{i1}, \mathbf{x_i}, \alpha_{2i}) = exp(x_{it}\beta_2 + \gamma_2 P_{it-1} + a_{20} + \bar{x}_i a_{21} + \alpha_{2i}) \tag{3.7}$$

$$P(L_{it}^{W/O} = 1|L_{it-1}^{W/O}, ..., L_{i1}^{W/O}, \mathbf{x_i}, \alpha_{3i}) = \Phi(x_{it}\beta_3 + \gamma_3 L_{it-1}^{W/O} + a_{30} + \bar{x}_i a_{31} + \alpha_{3i}) \tag{3.8}$$

Correlation between individual-specific errors $v_{it} = \alpha_i + u_{it}$ in any two periods is: $\lambda = corr(v_{it}, v_{is}) = \sigma_\alpha^2/(\sigma_\alpha^2 + \sigma_u^2)$ for $t, s = 2, ..., T; t \neq s$, where σ_u^2 is set to 1 in probit random effects model.

According to Heckman (1981*b*), two assumptions are typically invoked in estimating a dynamic random-effects model with binary outcomes. The first assumption is that the initial conditions are exogenous, or the start of the process coincides with the start of the survey period for each sampled individuals. The second assumption is that marginal probabilities are time-invariant. The first assumption is invalid when the error process is not serially independent (which is inevitable in the presence of the unobserved individual-effects), and the first observation is not the true initial outcome of the process, or the data is not collected at the beginning of the process. If the initial conditions are correlated with α_i, as is the case in the our data, a standard dynamic random-effects Probit model would overstate the level of state dependence (Heckman, 1981*b*; Stewart, 2007; Arulampalam and Stewart, 2009).

3.4.2.1 Wooldridge's Conditional ML Estimator

To address the initial condition problem, Wooldridge (2005) proposes using a conditional maximum likelihood estimator to model the joint distribution of $y_2, ...y_T$ conditional on the endogenous initial value y_1 and the strictly exogenous explanatory variables. We specify a model for α_i on y_1, instead of a model for y_1 on α_i which is Heckman's solution to initial conditions. In the case of the random effects Probit model, the individual-specific effect α_i is specified as below to account for the correlation between y_{i1} and α_i:

$$\alpha_i = a_0 + a_1 y_{i1} + \xi_i \tag{3.9}$$

Variables in x_i that are correlated with α_i have already been accounted for in the Mundlak-Chamberlain specification above. The new individual-specific random effect ξ_i is uncorrelated with y_{i1}. Substituting the equation above into the depression probability model gives us the following specification:

$$P(M_{it} = 1 | M_{it-1}, ..., M_{i1}, \mathbf{x_i}, \xi_{1i}) = \Phi(x_{it}\beta_1 + \gamma_1 M_{it-1} + a_{10} + \bar{x}_i a_{11} + M_{i1} a_{12} + \xi_{1i})$$
$$(i = 1, ..., N; t = 2, ..., T) \tag{3.10}$$

The likelihood function is given by:

$$L^M = \prod_{i=1}^{N} \int \{ \prod_{t=2}^{T} \Phi[(x_{it}\beta_1 + \gamma_1 M_{it-1} + a_{10} + \bar{x}_i a_{11} + M_{i1} a_{12} + \xi_{1i})(2M_{it} - 1)] \} g(\xi_{1i}) d\xi_{1i} \tag{3.11}$$

where $g(\xi_{1i})$ is the normal probability density function of the new individual-specific random effect ξ_{1i}. In the case where α_{1i} is normally distributed, the integral in the equation above can be evaluated using Gaussian–Hermite quadrature (Butler and Moffitt, 1982; Stewart, 2007; Arulampalam and Stewart, 2009). This dynamic, random effects Probit model can also be used to estimate the probability of being in different labour force status. Interacting the initial values of the dependent variable with time dummies would allow us to account for the correlation between the initial condition error and the errors in the other periods (Stewart, 2007; Arulampalam and Stewart, 2009).

For count variable of the number of physical health problems P_{it}, given $(P_{it-1}, ..., P_{i1}, \mathbf{x_i}, \xi_{2i})$, it follows a Poisson distribution with mean:

$$E(P_{it} | P_{it-1}, ..., P_{i1}, \mathbf{x_i}, \xi_{2i}) = exp(x_{it}\beta_2 + \gamma_2 P_{it-1} + a_{20} + \bar{x}_i a_{21} + P_{i1} a_{22} + \xi_{2i}) \tag{3.12}$$

where P_{i1} is observation of P_{it} in the first-period, ξ_{2i} is the new individual-specific random effect, and other variables are specified in the same way as above. Let m_{it}

denote the mean function of physical health P_{it} in equation (3.12), the density is given by:

$$\prod_{t=2}^{T} exp(-m_{it})(m_{it})^{P_t}/P_t! = (\prod_{t=2}^{T} m_{it}^{P_t}/P_t!)exp(\sum_{t=2}^{T} m_{it})d_i^{P_1+...+P_t} \tag{3.13}$$

After conducting both an empirical illustration and a set of simulation experiments, Arulampalam and Stewart (2009) find that neither Heckman nor Wooldridge estimators dominates the other in all cases, and both display satisfactory performance. Therefore, we report only Wooldridge's estimator.

In the case where attrition is an issue, Wooldridge's approach allows attrition to depend on the initial condition in an arbitrary way. Nonetheless, the comparison of summary statistics using balanced and unbalanced sample (Table 3.6) does not support a health-related attrition as the unbalanced sample we use reports similar risk of depression and worse physical health than the balanced sample.

3.4.3 Non-normality

Both Heckman and Wooldridge estimators might be sensitive to the assumption of normal individual effects (Stewart, 2007). Even under non-normality, ML is consistent and asymptotically normal under fixed T when N tends to infinity (Moral-Benito, 2013), but no longer efficient. It means that the reported standard errors are not consistent estimates of the true standard errors, and the p-values would be incorrect. Nonetheless, the degree of bias might not be serious. Simulations by Moral-Benito, Allison and Williams (2019) show that the ML estimator performs quite well compared with the GMM estimator in finite samples under both normal and non-normal data.

There are also approaches that we can use to adjust for standard errors under non-normality. Quasi–maximum likelihood (QML) relaxes the normality assumption and estimates the robust standard errors (Moral-Benito, Allison and Williams, 2019). Therefore, we report estimates with robust standard errors clustering in individual level.

3.4.4 Trivariate Dynamic Non-linear Random-Effects Models

To study cross- effects between depression, physical health and labour supply status, we extend the dynamic non-linear random-effects model in the previous section to include the one-period lags of other dependent variables. To account for the potential endogeneity of the lags of other dependent variables, or omitted variables that affect both health and labour supply decisions, we allow random effects on the three endogenous variables to correlate with each other. For example, an unobserved, individual propensity of experiencing health problems ξ_{1i} or ξ_{2i}, or idiosyncratic shocks

u_{1it} or u_{2it} affecting health, may be correlated with individual heterogeneity (time invariant) $\xi_{3i}^{W/O}$ or idiosyncratic shocks (time-varying) $u_{3it}^{W/O}$ affecting labour supply decision.

While the correlation of random intercepts suggests a contemporary relationship associated with omitted variables, estimates of the cross-lagged variables provide evidence for sequential causality. Therefore, the health effect on labour supply can come in the form of either contemporary health shocks or other unobservables (when correlation of random intercepts is significant), or declines in health (if the lagged health effect is significant). The distinction between the two effects is important for explaining the directions of effects between physical and mental health, and labour supply.

We estimate two trivariate models, using one of the two specifications of labour supply status in each of the trivariate models: doing agricultural work versus doing non-agricultural work or L_{it}^{W} for those currently in the labour force; being economically inactive (excluding the unemployed and those having never worked before) versus working (either doing agricultural or non-agricultural work), or L_{it}^{O} for the full sample. The lagged effects of labour force status in the depression and physical health equations are estimated by including an indicator of doing agricultural-work in the previous survey period (L_{it-1}^{F}), and an indicator of doing non-agricultural work in the previous survey period (L_{it-1}^{NF}). This means that those not working in the previous period (L_{it-1}^{O}) are used as the reference group, and the estimates of the lagged employment effects on health are relative to the lagged effect of not working. The models are specified as:

$$\begin{aligned} P(M_{it}=1|M_{it-1},...,M_{i1},P_{it-1},L_{it-1}^{F},L_{it-1}^{NF},\mathbf{x_i},\xi_{1i}) = \Phi(x_{it}\beta_1+\gamma_{11}M_{it-1} \\ +\gamma_{12}P_{it-1}+\gamma_{13}L_{it-1}^{F}+\gamma_{14}L_{it-1}^{NF}+a_{10}+\bar{x}_i a_{11}+M_{i1}a_{12}+\xi_{1i}) \end{aligned} \tag{3.14}$$

$$\begin{aligned} E(P_{it}|P_{it-1},...,P_{i1},M_{it-1},L_{it-1}^{F},L_{it-1}^{NF},\mathbf{x_i},\xi_{2i}) = exp(x_{it}\beta_2+\gamma_{21}M_{it-1} \\ +\gamma_{22}P_{it-1}+\gamma_{23}L_{it-1}^{F}+\gamma_{24}L_{it-1}^{NF}+a_{20}+\bar{x}_i a_{21}+P_{i1}a_{22}+\xi_{2i}) \end{aligned} \tag{3.15}$$

$$\begin{aligned} P(L_{it}^{W}=1|L_{it-1}^{W},...,L_{i1}^{W},M_{it-1},P_{it-1},\mathbf{x_i},\xi_{3i}^{W}) = \Phi(x_{it}\beta_1^{W}+\gamma_{31}^{W}M_{it-1} \\ +\gamma_{32}^{W}P_{it-1}+\gamma_{33}^{W}L_{it-1}^{W}+a_{30}^{W}+\bar{x}_i a_{31}^{W}+L_{i1}^{W}a_{32}^{W}+\xi_{3i}^{W}) \\ or \\ P(L_{it}^{O}=1|L_{it-1}^{O},...,L_{i1}^{O},M_{it-1},P_{it-1},\mathbf{x_i},\xi_{3i}^{O}) = \Phi(x_{it}\beta_1^{O}+\gamma_{31}^{O}M_{it-1} \\ +\gamma_{32}^{O}P_{it-1}+\gamma_{33}^{O}L_{it-1}^{O}+a_{30}^{O}+\bar{x}_i a_{31}^{O}+L_{i1}^{O}a_{32}^{O}+\xi_{3i}^{O}) \end{aligned} \tag{3.16}$$

Where $\bar{x}_i$ is the individual-specific average value of x_{it} over the sample period to account for the potential correlation between the unobserved individual heterogeneity and the observed exogenous variables, following Mundlak-Chamberlain approach.

Idiosyncratic errors $(u_{1it}, u_{2it}, u_{3it}^{W/O})$ are not explicitly specified in the probability equations as they have been used to form the probit and poisson models. They are assumed to be independent over time and jointly normally distributed with unit variances. The random effects $(\xi_{1i}, \xi_{2i}, \xi_{3i}^{W/O})$ are assumed to be jointly normally distributed with variances of σ_1^2 and σ_2^2, $\sigma_3^{2(W/O)}$, respectively. The pair-wise correlations between random effects are ρ_{12}, $\rho_{23}^{W/O}$ and $\rho_{13}^{W/O}$. If $\rho_{12} = \rho_{23}^{W/O} = \rho_{13}^{W/O} = 0$, the cross-lagged terms are weakly exogenous and the three equations can be estimated on their own. The likelihood function of the trivariate model is given by:

$$L = \prod_{i=1}^{N} \int_{\xi_3^{W/O}} \int_{\xi_2} \int_{\xi_1} P_i(\xi_1, \xi_2, \xi_3^{W/O}) f_i(\xi_1, \xi_2, \xi_3^{W/O}) d\xi_1 d\xi_2 d\xi_3^{W/O} \qquad (3.17)$$

where f_i is the trivariate, joint normal density of $(\xi_1, \xi_2, \xi_3^{W/O})$ for individual i, and P_i is the joint probability of observing the trivariate sequence for individual i as a function of the random effects .

Identification of the structural model relies on the assumption of tivariate normality of the individual random effects. Initial conditions of the dependent variables are included as the exclusive variables to improve identification of the model, apart from different distributions of the endogenous variables.

3.5 Empirical Results

In this section, we report and compare the results for the various model specifications above. Table 3.7 to 3.10 compare the estimates of different models for outcomes of the risk of depression risk (M_{it}), the number of physical health problems (P_{it}), the risk of non-working as opposed to working (either taking agricultural or non-agricultural jobs, L_{it}^O), and the risk of doing agricultural work as opposed to non-agricultural work (L_{it}^W), respectively.

Both state dependence, observable individual characteristics and unobserved individual heterogeneity can explain persistence in individual health and labour supply outcomes. We will explain by comparing different models how the trivariate model enable us to distinguish between the three factors and estimate the real state dependence that other simpler models fail to do.

We start with estimating a single dynamic random effects model (columns 1 and 5) where the endogenous, cross-effects of health and labour supply are not included. This would bias the estimates of state dependence and other covariates in the model due to omitting variables. We develop the single dynamic random-effects models by including as extra explanatory variables the one-period lags, or cross-effects of other health outcomes and labour force status, as shown in column 3 and 7.

Although the cross-effects are statistically significant in explaining the current health and labour supply status, they are nonetheless endogenous. Health and labour supply behaviour are likely to be correlated with each other due to their strong correlations with other omitted individual socio-economic status or community variables, preferences for health that determine both current health and current labour supply status. Health and labour supply decisions are also directly correlated.

Column 4 and 8 report estimates of the trivariate models that deal with the endogeneity problem of the cross-effects. Apart from coefficients, we also report pair-wise correlations of random effects. If $\rho_{13W/O} = 0$ and the estimated effects of L^F_{it-1} or L^{NF}_{it-1} are significantly different from zero in predicting M_{it}, then L^F_{it-1} or L^{NF}_{it-1} is weakly exogenous in the model of M_{it} (Stewart, 2007). We run two trivariate models using one of the two different specifications of labour force status, and separately for men and women. Because estimation of the depression equation and physical health equation changes only slightly in the trivariate models that use different specifications for the labour supply equation, we report estimation from the trivariate model that estimates the joint probability of depression, number of physical problems, and probability of non-working.

To simplify the table, we do not report the coefficients of the individual-level means variables, location dummies and age-group dummies. Table 3.7 to 3.10, column 1 and 5 report the estimates of the single dynamic, random-effects models (Equation 3.10 and 3.12), column 2 and 5 report the estimates of the dynamic, pooled models with cross-effects, column 3 and 6 report the estimates of the dynamic, random-effects models with cross-effects, and column 4 and 8 report the estimates of the trivariate dynamic, random-effects models accounting for both cross effects and correlations between random effects of different outcomes. At the bottom of the tables, we report estimated variances of the random effects, and pair-wise correlations of random effects. Column 1-4 and column 5-6 report estimates for men and women, respectively.

3.5.1 Individual Heterogeneity

We estimate unobserved heterogeneity explicitly by including Gaussian random effects in the models. Coefficients of individual random effects ξ_i are set to 1 for identification of their variances (σ^2_ξ). The intra-class correlation coefficient, calculated as $\sigma^2_\xi/(1+\sigma^2_\xi)$, measures how much the latent error variance can be explained by the individual random effects. It is reported for single dynamic random effects models.

Adding the significant cross effects of the lagged health and employment variables sees the intra-class correlation coefficient decline, indicating that part of the unobservable heterogeneity in the single dynamic model in column 1 and 5 can be explained by the omission of the cross-effects of health and employment outcomes. Specifically, the intra-class correlation coefficient declines from 36.2% (column 1) to

30.5% (column 3) in explaining the probability of depression for men and from 34% (column 5) to 28.5% (column 7) for women, after we account for the significant cross effects of the lagged physical health (P_{it-1}) and the lagged employment variables (L^{F}_{it-1} L^{NF}_{it-1}). This suggests that the past physical health conditions and labour supply status contribute to part of the omitted effects that determine the current depression for men. Nonetheless, the past physical health or labour supply status might be correlated with current depression due to the fact that they are closely related to the same sets of the socio-economic status variables or community variable, macroeconomic variables etc. that are omitted from the regression. By allowing for the random effects of the health and labour supply variables to correlate with each other, we can identify the direct effect of past health or labour supply on the current risk of depression. This however, increases the level of the intra-class correlation coefficient in the trivariate random effects model in column 4 and 8, compared to its level in column 3 and 7. Nonetheless, levels of the intra-class correlation coefficient in the trivariate models are still higher than its levels in models that do not account for the cross effects (column 1 and 5). Magnitudes and the significant levels of the cross-effects also decline in the trivaraite models, due to the positive and significant correlations that both physical health and depression, or both non-working and depression have with the same set of SES omitted from the regression ($\rho_{12} > 0$, $\rho_{13^O} > 0$). This means that cross-effects in the single RE models of depression not accounting for the endogeneity of the cross-effects are over-estimated because they catch up the effects of other omitted variables that predict both higher risk of depression and more physical health problems, or both higher risk of depression and risk of non-working. Based on the intra-class correlation coefficient in trivariate models, the significant effect of past physical health conditions explain more than 4% of the omitted effects that determine the current depression for both men and more than 2% for women.

Column 1 in Table 3.8 shows that individual heterogeneity explains 19.2% for the number of physical health problems for men, and declines slightly to 18.9% after including the significant effect of past depression and the non-significant effect of past employment. The effect of past depression becomes insignificant in the trivariate models after we account for the fact that physical problems and depression are positively related to the same sets of omitted effects. Variance of the random effect of physical health problems and the intra-class correlation coefficient changes negligibly across different models, due to the overall insignificant cross-effects. The changes in the intra-class correlation coefficient across different models are also minor for women.

As for the risk of non-working, individual heterogeneity declines from 40% to 36.1% for men after including the significant cross-effects of depression and physical health, and declines from 27.4% to 23.5% for women, as shown in Table 3.9. It suggests that the effects of past depression and past physical problems are caught up by the lagged labour supply variable or the SES variables in models that do not account for the cross-effects (column 1 and 5). The cross effects change in magnitudes but are still significant in the trivariate models, and the intra-class correlation coefficient increases slightly but are still lower than its levels in the models not controlling for cross-effects.

The cross-effects of health are not significant in modelling the risk of taking agricultural versus non-agricultural jobs for either men or women. As a result, the levels of intra-class correlation coefficient do not change substantially across different models.

Compared with other outcomes, the relatively low magnitudes of intra-class correlation coefficients and variances of the random effects for the number of physical health problems indicate that physical health can be better explained more by the persistence of diseases or functional limitations themselves, especially for women. The individual unobservables that play an important role in explaining depression risk and labour supply risk can be life events or shocks that cause short-term depression, or other factors like financial stress, community-level characteristics that affect both health and labour supply decisions such as local labour market conditions, public health infrastructure.

To shed a light on how random-effects models can distinguish between state dependence and the individual unobservables that pooled models fail to, we compare with the estimates of pooled models, which are reported in column 2 and 6 within Table 3.7 to 3.10. The lagged dependent variable, or state dependence, is upwardly biased with a higher level significance in the pooled model because it catches both state dependence and individual unobservables. The cross effects can be either over- or under-estimated in the pooled model, due to their correlations with the omitted effects. The direction of the bias depends on the signs of the pair-wise correlations between random effects of the two endogenous variables. Nonetheless, we cannot compare directly the magnitudes of the estimates in the pooled models with those in the RE models due to different scaling of the error variances. We can compare instead the significant levels, and the relative effects of pairs of variables.

3.5.2 State Dependence

We find strong and significant state dependence for both depression, physical health problems and labour supply status, with gender-specific variation in the magnitudes. This suggests that health deterioration among older, informal workers are persistent and the financial risks associated with health expenditures and loss of labour income will be long-daunting. More generous health interventions and modest social pension programmes can make persistent, positive differences on the older, informal workers' health conditions and labour productivity.

state dependence in the single dynamic RE models (columns 1 and 5) declines slightly in magnitudes after we include the cross-effects, due to its correlation with the newly included, endogenous, cross-effects of health and labour supply.

Pooled models significantly over-estimate the state dependence because the lagged dependent variable in the pooled models catches both state dependence and individual unobservables. If we compare the random effects specifications (column 3 and 7) to the pooled models, the p-values on state dependence are smaller and p-values on the initial health states are larger. The magnitudes of the state dependence relative to those of the initial conditions are reversed in the random effects models, where state dependence is smaller than the initial conditions, compared to the pooled models. Similar findings are reported in Contoyannis, Jones and Rice (2004)'s study that compares dynamic ordered Probit models with pooled models.

Accounting for the fact that cross-effects are correlated with the random effects in the trivariate models again changes the levels of state dependence. Specifically to the different health and labour supply outcomes, state dependence is significant in predicting depression for women but is only marginally significant for men (Table 3.7), and the level halves for men. Recall that women overall report a higher rate of depression. The results suggest that the higher persistence of depression among women might contribute to the gender-specific inequality in the depression risk. State dependence is highly significant for physical health problems and has similar magnitudes for men and women (Table 3.8). The result suggests a long-term nature of physical health problems, typical for chronic diseases. Physical health is affected more by long-term factors like health behaviours, living conditions, physical activities etc., while depression can be a result of unexpected shocks like losing a close family members, financial stress, job loss.

Women show a higher level of persistence in non-working than men do, meaning that women are less likely to be back to the labour force after they stop working. The state dependence in non-working has been explained by high searching costs, stigma effects, human capital depreciation among the economically inactive in the literature (Hyslop, 1999), and they might explain the gender difference in the likelihood of returning to work after being economically inactive. Nonetheless, the economically inactive tend to be in old ages (on average over 65) and report almost twice as many physical health problems as the current workers do. They might choose not to be back to work instead of couldn't find one.

Men and women show similarly higher level of state dependence in the probability of doing agricultural work or non-agricultural work. This means that the likelihood of a transit between the two sectors are similar for both gender groups, although women in our sample are less likely to work in the non-agricultural sectors than men do, as shown in Table 3.2.

Table 3.7: Comparison of Models for Depression Risk

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE	Pooled	RE	Tri	RE	Pooled	RE	Tri
M_{it-1}	0.041	0.435***	0.070	0.048	0.118	0.562***	0.148*	0.115
	(0.098)	(0.049)	(0.098)	(0.096)	(0.078)	(0.039)	(0.081)	(0.078)
M_{i1}	0.992***	0.369***	0.739***	0.763***	0.966***	0.322***	0.769***	0.799***
	(0.112)	(0.048)	(0.104)	(0.102)	(0.092)	(0.039)	(0.091)	(0.088)
P_{it-1}		0.074***	0.087***	0.067***		0.063***	0.075***	0.049***
		(0.007)	(0.009)	(0.009)		(0.005)	(0.007)	(0.007)
L^F_{it-1}		-0.081	-0.118*	-0.105		-0.003	-0.021	-0.042
		(0.051)	(0.062)	(0.066)		(0.038)	(0.047)	(0.051)
L^{NF}_{it-1}		-0.216***	-0.277***	-0.265***		0.054	0.059	0.050
		(0.067)	(0.082)	(0.086)		(0.060)	(0.075)	(0.079)
rrural2	0.327***	0.251***	0.312***	0.316***	0.179**	0.110*	0.151*	0.153*
	(0.104)	(0.081)	(0.101)	(0.102)	(0.084)	(0.066)	(0.084)	(0.086)
rage	0.087	0.071	0.083	0.071	0.118**	0.047	0.071	0.065
	(0.062)	(0.050)	(0.060)	(0.062)	(0.047)	(0.039)	(0.048)	(0.049)
rage_sq	-0.071	-0.062*	-0.073*	-0.063	-0.099***	-0.047*	-0.068**	-0.063*
	(0.044)	(0.036)	(0.043)	(0.044)	(0.034)	(0.028)	(0.034)	(0.035)
rprimary	0.025	0.035	0.039	0.040	-0.124**	-0.068	-0.091	-0.099*
	(0.060)	(0.045)	(0.056)	(0.057)	(0.057)	(0.042)	(0.056)	(0.058)
rsecondary	-0.128*	-0.084	-0.105	-0.116*	-0.177***	-0.123**	-0.164**	-0.174**
	(0.069)	(0.052)	(0.065)	(0.066)	(0.068)	(0.051)	(0.067)	(0.070)
rhighabove	-0.190*	-0.060	-0.082	-0.098	-0.315***	-0.199**	-0.258**	-0.274**
	(0.101)	(0.076)	(0.095)	(0.097)	(0.114)	(0.085)	(0.112)	(0.118)
rmarried	-0.557**	-0.521***	-0.557***	-0.583***	-0.759***	-0.612***	-0.745***	-0.785***
	(0.219)	(0.188)	(0.215)	(0.218)	(0.160)	(0.133)	(0.158)	(0.161)
hchild	0.023	0.035	0.035	0.034	-0.039	-0.034	-0.054	-0.055
	(0.049)	(0.042)	(0.048)	(0.049)	(0.043)	(0.036)	(0.043)	(0.044)
hhhres	-0.015	-0.015	-0.017	-0.019	-0.001	0.004	0.005	0.005
	(0.021)	(0.018)	(0.020)	(0.021)	(0.016)	(0.014)	(0.016)	(0.017)
lnhhadurbl	-0.006	-0.007	-0.008	-0.006	0.008	0.001	0.003	0.003
	(0.015)	(0.013)	(0.015)	(0.015)	(0.012)	(0.010)	(0.012)	(0.012)
y2015	0.151***	0.131***	0.124***	0.132***	0.175***	0.160***	0.167***	0.181***
	(0.042)	(0.037)	(0.042)	(0.042)	(0.034)	(0.030)	(0.035)	(0.035)
urban_nbs	0.008	-0.086	-0.104	-0.088	-0.161	-0.174**	-0.226**	-0.256**
	(0.157)	(0.123)	(0.152)	(0.155)	(0.112)	(0.088)	(0.115)	(0.120)
ξ_i	1.000		1.000	1.000	1.000		1.000	1.000
	.		.	.	.		.	.
_cons	-3.152	-2.720*	-3.047	-2.612	-3.427**	-1.574	-2.148	-1.803
	(1.979)	(1.599)	(1.922)	(1.973)	(1.491)	(1.223)	(1.508)	(1.547)
Provinces	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AgeGroups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{x}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
σ^2_1	0.638***		0.428***	0.491***	0.696***		0.530***	0.635***
	(0.141)		(0.118)	(0.122)	(0.119)		(0.109)	(0.115)
ICC	0.389***		0.300***	0.329***	0.410***		0.346***	0.388***
	(0.086)		(0.134)		(0.070)		(0.122)	
ρ_{12}				0.559***				0.604***
				(0.061)				(0.050)
ρ_{13^O}				0.163**				-0.030
				(0.081)				(0.066)
ρ_{13^W}				0.228**				-0.084
				(0.105)				(0.251)
Observation	7015	6777	6777	6834	9620	9035	9035	9106
Log Likelihood	-3654	-3483	-3472	-19019	-5674	-5292	-5271	-28641

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. ICC stands for the intra-class correlation coefficient and is calculated as $\sigma^2_\xi/(1+\sigma^2_\xi)$. Other covariates not reported include the individual-level means for the time-varying variables among them, demographic indicators of 7 types of administrative areas and of 28 provinces. Sample of individuals aged 40 or above at the first wave, staying for wave 1 and 2 and not missing in the outcome of M_{it} in column 1 and 5, or not missing in outcomes of M_{it}, P_{it} and L^O_{it} for the rest of models, not belonging to the group of formal retirees or formal workers reaching the compulsory retirement ages, are selected for the estimation.

Table 3.8: Comparison of Models for Physical Health Problems

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE	Pooled	RE	Tri	RE	Pooled	RE	Tri
P_{it-1}	0.011**	0.100***	0.013***	0.016***	0.015***	0.078***	0.012***	0.015***
	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.003)	(0.004)	(0.004)
P_{i1}	0.182***	0.064***	0.178***	0.174***	0.138***	0.060***	0.138***	0.135***
	(0.006)	(0.005)	(0.007)	(0.007)	(0.004)	(0.004)	(0.005)	(0.005)
M_{it-1}		0.117***	0.083***	0.014		0.123***	0.108***	0.034**
		(0.020)	(0.018)	(0.019)		(0.014)	(0.013)	(0.014)
L^F_{it-1}		-0.038	-0.041	0.002		-0.050***	-0.046***	-0.030*
		(0.025)	(0.025)	(0.026)		(0.015)	(0.016)	(0.017)
L^{NF}_{it-1}		-0.066**	-0.045	-0.003		-0.078***	-0.066**	-0.053*
		(0.033)	(0.033)	(0.033)		(0.029)	(0.028)	(0.028)
rrural2	-0.047	-0.010	-0.028	-0.028	-0.057**	-0.039	-0.052*	-0.051*
	(0.038)	(0.040)	(0.041)	(0.041)	(0.028)	(0.027)	(0.031)	(0.030)
rage	0.036	0.052**	0.039	0.037	0.026*	0.040**	0.033*	0.038**
	(0.024)	(0.025)	(0.026)	(0.026)	(0.015)	(0.017)	(0.017)	(0.017)
rage_sq	-0.020	-0.033*	-0.024	-0.021	-0.015	-0.026**	-0.021*	-0.025**
	(0.017)	(0.018)	(0.019)	(0.019)	(0.011)	(0.012)	(0.012)	(0.012)
rprimary	-0.005	-0.001	0.009	0.006	-0.033	-0.032*	-0.035	-0.038*
	(0.026)	(0.022)	(0.027)	(0.027)	(0.021)	(0.018)	(0.021)	(0.022)
rsecondary	-0.048	-0.038	-0.049	-0.053	-0.080***	-0.068***	-0.077***	-0.086***
	(0.031)	(0.027)	(0.032)	(0.032)	(0.027)	(0.024)	(0.028)	(0.028)
rhighabove	-0.137***	-0.103**	-0.126***	-0.130***	-0.099**	-0.054	-0.073	-0.077
	(0.044)	(0.041)	(0.048)	(0.047)	(0.046)	(0.041)	(0.050)	(0.050)
rmarried	-0.144**	-0.136	-0.143*	-0.122	-0.063	-0.055	-0.063	-0.058
	(0.070)	(0.087)	(0.079)	(0.079)	(0.044)	(0.053)	(0.050)	(0.050)
hchild	-0.003	-0.018	-0.011	-0.012	-0.012	-0.021	-0.020	-0.022*
	(0.016)	(0.021)	(0.017)	(0.017)	(0.013)	(0.015)	(0.013)	(0.013)
hhhres	0.003	0.007	0.008	0.008	0.012***	0.016***	0.016***	0.016***
	(0.006)	(0.008)	(0.006)	(0.006)	(0.004)	(0.006)	(0.005)	(0.005)
lnhhadurbl	-0.007	-0.010*	-0.010**	-0.010*	-0.007**	-0.005	-0.004	-0.004
	(0.005)	(0.006)	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.004)
y2015	0.192***	0.184***	0.206***	0.201***	0.211***	0.208***	0.222***	0.216***
	(0.013)	(0.018)	(0.014)	(0.014)	(0.010)	(0.013)	(0.011)	(0.011)
urban_nbs	0.069	0.075	0.101	0.104	-0.030	-0.074*	-0.087*	-0.090*
	(0.060)	(0.061)	(0.068)	(0.068)	(0.041)	(0.040)	(0.046)	(0.047)
ξ_i	1.000		1.000	1.000	1.000		1.000	1.000
	.		.	.	.		.	.
_cons	-0.875	-1.213	-1.003	-0.891	-0.122	-0.483	-0.340	-0.432
	(0.761)	(0.798)	(0.837)	(0.844)	(0.491)	(0.528)	(0.544)	(0.544)
Provinces	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AgeGroups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{x_i}$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
σ^2_2	0.231***		0.214***	0.216***	0.153***		0.136***	0.140***
	(0.013)		(0.013)	(0.013)	(0.008)		(0.008)	(0.013)
ICC	0.188***		0.176***	0.178***	0.133***		0.120***	0.123***
ρ_{12}				0.559***				0.604***
				(0.061)				(0.050)
ρ_{23^O}				0.360***				0.184***
				(0.066)				(0.058)
ρ_{23^W}				0.036				-0.142
				(0.084)				(0.236)
Observation	7874	6821	6821	6834	10376	9081	9081	9106
Log Likelihood	-15488	-13804	-13310	-19019	-22494	-19897	-19460	-28641

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Sample of individuals aged 40 or above at the first wave, staying for wave 1 and 2 and not missing in the outcome of P_{it} in column 1 and 5, or not missing in the outcomes of M_{it}, P_{it} and L^O_{it} for the rest of models, not belonging to the group of formal retirees or formal workers reaching the compulsory retirement ages, are selected for the estimation.

3.5.3 Dynamic Cross-Effects of Health and Employment

In this section we will discuss about the estimated cross-effects between depression, physical health problems and labour supply status. We will compare between the dynamic cross-effects pooled models, the single dynamic cross-effects RE models, and the trivariate models. We also interpret the correlations between the unobserved heterogeneity that predicts different health and labour supply outcomes, which are estimated in the trivariate models.

3.5.3.1 Depression and Physical Health Problems

Physical health declines significantly increase the risk of depression in the subsequent period for both men and women across all three models. The effect for women is less than 2/3 the effect for men, meaning physical problems play a less important role in explaining women's depression than in men's.

The effect of past physical health (P_{it-1}) is more significant in the pooled model than it is in the trivariate RE models, because it picks up the effects of the unobserved variables that significantly predict both depression and more physical health problems. The unobserved effects are accounted for by the significant and positive correlation between the two random effects ($\rho_{12} > 0$) in the trivariate models. ρ_{12} is slightly larger for men than it is for women, so the bias is larger for men than for women. If we compare the estimates of the dynamic RE models accounting for cross-effects with those of the trivariate models, we can see that magnitudes of the cross-effect of past physical health change after we account for the correlation between the two random effects (ρ_{12}). Specifically, the positive effect of past physical problems declines in magnitudes by about 29.4% for men and 43.5% for women. The reason, as mentioned above, is that part of the cross-effects of depression and physical health in the single RE models comes from the individual unobservables that determine both processes. They can be health behaviours, lifestyle choices and social capital (Ohrnberger, Fichera and Sutton, 2017*b*), medical complications or stress in patients having certain chronic diseases (Katon, Lin and Kroenke, 2007), or genetic factors that increase risks of both depressive syndromes and other chronic diseases (Bruce, 2001; Taylor, Aizenstein and Alexopoulos, 2013). By estimating a trivariate model, we condition on the initial conditions of physical health and labour supply status in estimating the cross-effects of them on current depression.

As for the effect of depression on the subsequent physical health (Table 3.8), it is not significant for men in the trivariate model, but are significant and much higher in magnitude for women. It means that depression plays a more vital role in explaining women's physical health declines. The results also suggest that the interrelated effects between depression and physical health are not symmetric and physical problems have a stronger impact on the subsequent mental health than the other way round.

In summary, our empirical evidence supports the bidirectional effects between depression or anxiety and severity of medical illness in medical and psychological studies. We do not study how they affect each other, either by changing health behaviours or biological complications, or genetic factors that we mention in the conceptual framework.

3.5.3.2 Depression and Labour Supply

Estimates of the single dynamic cross-effects RE model in Table 3.7 show that working in the non-agricultural sector (L^{NF}_{it-1}) significantly reduces the subsequent risk of depression for men, compared with the economically inactive (the reference group). The effect more than doubles for non-agricultural workers compared with the agricultural workers (L^{F}_{it-1}), who show no difference in the risk of depression compared with the non-workers in trivaraite models. This is consistent with the literature that finds working, especially non-labour-intensive work, to have a preservative effect on cognitive functioning and a healthy mental condition (Llena-Nozal, Lindeboom and Portrait, 2004). Nonetheless, the effect of employment is not significant for women in our sample. It indicates that risks factors for depression are different for men and women, with labour supply behaviour playing a more important role in maintaining men's mental health condition.

The positive effect of non-agricultural work on mental health (L^{NF}_{it-1}) for men is over-estimated in the single pooled and RE models in Table 3.7, as they do not account for the omitted effects that predict both non-working and higher risk of depression ($\rho_{13^O} > 0$), or the omitted effects that predicts doing agricultural work and a higher risk of depression ($\rho_{13^W} > 0$). In other words, the non-agricultural workers in our sample are in better mental health conditions and not accounting for the selection bias can over-estimate the negative coefficients of doing non-agricultural on the risk of depression. The omitted effects are not significant for women, and the effect of labour supply does not change in single pooled or RE models for women. We can say that the lagged employment variables are weakly exogenous to women's, but endogenous to men's risk of depression.

As for the effect of depression on labour supply decisions, estimates of the single dynamic cross-effects RE model in Table 3.9 show that depression significantly increases the probability of re-entering the labour market in the subsequent period for men. The effect becomes only marginally significant and halves in magnitude for women. In the trivaraite model, the effect becomes more significant and increases in magnitude for men, but becomes insignificant for women. The correlation between the random effect of depression and that of the probability of not working, ρ_{13^O}, is also positive and significant for men but not for women. It suggests that not controlling for the unobserved effect that the economically inactive men are more likely to be depressed than men in the labour force will bias the cross-effect of depression downward in the single dynamic RE models and pooled models. The summary statistics (Table 3.2) in the previous section also shows that the economically inactive are in the lowest socio-economic status and suffer from the highest rate of depression compared with people in other labour force status.

The unobservables can be financial stress or loss of income and social capital in work environment, changes in health behaviour etc. that make the economically inactive men subject to higher risk of depression (Eibich, 2015). These same factors can push them to re-enter the labour market, which explains the higher probability of labour market participation for men following a period of depression. The unobservables are not significant for women, neither is the effect of past depression. This means that labour supply, or a quality job plays a more important role in maintaining good subjective well-being for men than for women. The reason can be that financial stress does not have an equally important impact on women's mental health and labour supply decisions because of the different family roles they take. Intra-household labour supply allocation and spousal effects have been found to play a role in explaining gender-specific differences in labour supply responses to health shocks. As the survey we use collect information of both respondents and their spouses (if present), by studying gender-specific pattern of health and labour market transitions, we also shed light on different family roles that men and women take in rural China with men being the main income-earners for low-socio-economic status households. Women in the rural China or in low socio-economic status households tend to have lower bargaining power and spend more time on within-household informal labour than men do. Finally, men tend to have higher employment attachment than women do, and their mental health conditions deteriorate if they become unemployed. As a result, they are more likely to be back to work (García-Gómez, Jones and Rice, 2010).

The finding is against the hypothesis that depression can discourage work because it increases costs of effort and lowers productivity in jobs (Baranov et al., 2020). Note that the hypothesis and its supporting, empirical evidence are often based on the context of formal employees in developed countries. Our results show that for the older, informal workers in a developing country or from low socio-economic status households, depression is associated with being economically inactive and predicts the subsequent re-entering into the labour market for men, but has no significant correlation with women's labour participation decisions. The cross-countries difference in responses to depression can be due to the different causes of depression, which tend to be financial stress, unemployment and the associated loss of incomes in our context.

Past depression does not affect the probability of choosing agricultural work over non-agricultural work for either men or women across all models, as shown in Table

3.10. The estimated correlation between random effects of depression and that of doing agricultural work (ρ_{13W}) in the trivariate models is positive and significant for men but not for women, suggesting that men doing agricultural jobs are subject to higher risks of depression than men doing non-agricultural jobs. The same correlation is found in the summary statistics (Table 3.2). The reasons might be that men doing agricultural work tend to work for longer hours and not have time for leisure activities and social interaction, and are more likely to be subject to financial stress. Regional differences in health care system and community characteristics might also play a role in explaining the sectoral difference in mental health, as agricultural workers almost all stay in the rural areas and non-agricultural workers are more likely to be coming from more urbanized areas that have higher proportion of non-agricultural sectors.

3.5.3.3 Physical Problems and Labour Supply

Labour market status does not seem to affect physical health in the subsequent period for men, as neither agricultural workers (L^{F}_{it-1}) nor non-agricultural workers (L^{NF}_{it-1}) report significantly less physical health problems than the economically inactive (the reference group) do in Table 3.8, except for the pooled model. It means that being economically inactive does not increase physical problems for men within a short period of time. Women working in the non-agricultural sector report less physical problems than the economically inactive women (the reference group) do, though the effect is only marginally significant in the trivariate model, after we account for the fact that the economically inactive women suffer from more physical health issues than women in the labour force do. Literature has also provided evidence on the negative effect of non-working on mental and physical health (Clark and Oswald, 1994), and the correlations between poor health and non-employment (Kalwij and Vermeulen, 2008; Haan and Myck, 2009). Our findings suggest that labour market policies and other supports that relieve the burden of informal labour supply within households and encourage formal employment for women in the low socio-economic status conditions can benefit their mental health and potentially physical health due to the significant cross-effects between the two outcomes for women.

As for the effect of physical health on individual labour supply outcomes, estimates of the single dynamic cross-effects RE model in Table 3.9 show that, physical health problems significantly increase the probability of not working in the subsequent period, higher for men than for women. The effect decreases in magnitude but remain significant in the trivariate models after we control for the significant and positive correlations between the random effects of physical health and of non-working (ρ_{23O}¿0). The correlation between the random effects almost halves but is still significant for women.

The adverse physical health effect increases by almost 44% for men compared to it is for women in the trivariate models. This means that women and especially men suffering from more physical health problems are more likely to stop working in the subsequent period, because men tend to take labour-intensive jobs that are more likely to be affected by physical limitations. The higher contemporary correlation between physical problems and non-working (ρ_{23O}) for men also supports the hypothesis. Women might stop working and take care of families before they are physically incapable of working. In other words, physical limitations are not so important for women as they are for men in determining the timing of leaving the labour force, and other determinants such as caring responsibility might play an important role in women's labour supply decisions.

The unobservables that affect both men and women's physical health and labour market participation can be community factors like prices of health care and health inputs, availability and quality of health care services, public health infrastructure, environmental factors, local labour market conditions and economic development. These factors can cause both lower level of public health and higher level of non-working Lei et al. (2014*b*). ρ_{23O} is more significant and larger in magnitude than ρ_{13O}, meaning that physical health is more closely related to labour participation decisions than depression or subjective well-being are for older, informal workers in our sample.

The physical health effect on labour supply can be sector-specific and affects productivity of agricultural workers or workers taking labour-intensive jobs more than non-agricultural workers. Individuals suffering from physical health deterioration might shift from labour-intensive agricultural sector to non-agricultural business (Adhvaryu and Nyshadham, 2017). Nonetheless, we fail to find any significant physical health effect on the probability of transition between agricultural and non-agricultural jobs for either men or women, as shown in the trivariate models in Table 3.10. This may suggest that there exists a barrier between agricultural and nonagricultural sector for older workers. Older workers may choose to reduce working hours instead of changing job sectors after health declines. Sectoral choice is usually made in the early career stage when physical health issues are not a big concern, while the decision to exit the labour force are made in the old ages when health becomes an important risk factor for the older workers.

We also study the effect of physical health deterioration on the risk of non-working separately for agricultural and non-agricultural workers in Table 3.11. Physical health decline sees more non-agricultural workers than agricultural workers stop working in the subsequent period. The negative labour supply effect more than doubles for female non-agricultural workers compared to female agricultural workers. This suggests that given the same level of physical health declines, agricultural workers are more likely to continue with working than non-agricultural workers do, and men more likely than women. The reason can be their different level of preparation for life without working and non-agricultural workers tend to have more savings for their retirement.

The health effects on labour supply we estimate are not likely to be subject to substantial regional differences, which will be a concern if there are substantial regional differences in the generosity of health care schemes and social security programmes[12]. The limited coverage of social safety net in the rural and less-developed areas of China where there are biggest proportions of informal workers means that the income effect of social security programme is unlikely to be significant in determining labour market entry and exit of the older, informal workers we study. Evidence on the labour supply effect of health insurance relies on the insurance being tied to employment, or the employment lock effect (French and Jones, 2011; Garthwaite, Gross and Notowidigdo, 2014). Nonetheless, health insurance is not a major determinant of the labour supply and welfare exit decisions of low income groups (Gruber and Madrian, 2002). This also applies to informal workers in China who are usually covered by the New Cooperative Medical Scheme that are less generous than employees' health insurance[13].

3.5.4 Socio-Economic Status

Finally, we study the effects of socio-economic status on health and labour supply outcomes and compare estimates across different models. We will focus on estimates from the trivariate model.

Rural hukou serves as a proxy for people's living environment that they are born into and grow up, as urban hukou holders cannot change into rural hukous. As shown in Table 3.7, being born in rural areas and holding a rural-hukou ($rrural2$) significantly predicts a higher risk of depression for both men and women, compared to their urban counterparts. The effect declines by 40% and is less significant for women compared to it is for men. Magnitudes and significant levels change negligibly after we include the cross-effects. Rural-hukou does not predict more physical problems for men or women (Table 3.8). Women holding rural-hukous are more likely to stay in the labour force than their urban counterparts, though the effect is only marginally significant (Table 3.9). Rural hukou is significantly related to taking agricultural jobs as opposed to non-agricultural jobs for both men and women (Table 3.10), with the effect almost double for women. It suggests that women born or raised in rural areas are less likely to take non-agricultural jobs than their male counterparts do when they grow up and enter the labour force. The literature has found that disadvantaged early life circumstances can have long-term, negative effects on educational achievement,

[12]The literature tends to agree that social insurance programmes that increase income for the non-workers, such as unemployment insurance (Krueger and Meyer, 2002) and disability insurance (Gruber, 2000; French and Song, 2014) have negative impacts on individual labour supply.

[13]Studies on welfare benefits of the New Cooperative Medical Scheme have found that it has no significant effect on out-of-pocket expenditures or health status measured by either self-reported health or sickness and injury (Lei and Lin, 2009), mainly due to the low generosity level of the scheme (Sun, 2020).

Table 3.9: Comparison of Models for Non-Working Risk

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE	Pooled	RE	Tri	RE	Pooled	RE	Tri
L^O_{it-1}	0.688***	1.057***	0.604***	0.625***	0.709***	0.919***	0.650***	0.650***
	(0.145)	(0.073)	(0.156)	(0.154)	(0.102)	(0.051)	(0.105)	(0.105)
L^O_{i1}	1.121***	0.418***	1.065***	1.051***	0.883***	0.486***	0.855***	0.857***
	(0.218)	(0.076)	(0.222)	(0.219)	(0.141)	(0.053)	(0.141)	(0.140)
M_{it-1}		-0.115**	-0.150**	-0.184***		-0.068*	-0.070*	-0.061
		(0.049)	(0.062)	(0.064)		(0.035)	(0.039)	(0.042)
P_{it-1}		0.076***	0.099***	0.086***		0.055***	0.063***	0.057***
		(0.008)	(0.012)	(0.012)		(0.006)	(0.007)	(0.007)
rrural2	-0.000	0.012	0.019	0.011	-0.156**	-0.110	-0.125	-0.127
	(0.109)	(0.092)	(0.119)	(0.118)	(0.077)	(0.070)	(0.081)	(0.081)
rage	0.022	-0.020	-0.052	-0.053	0.004	-0.010	-0.005	-0.006
	(0.075)	(0.066)	(0.084)	(0.083)	(0.050)	(0.048)	(0.055)	(0.055)
rage_sq	0.029	0.049	0.083	0.083	0.031	0.036	0.037	0.038
	(0.053)	(0.047)	(0.060)	(0.059)	(0.035)	(0.034)	(0.039)	(0.039)
rprimary	-0.038	0.027	0.025	0.024	-0.025	-0.004	-0.002	-0.003
	(0.064)	(0.052)	(0.069)	(0.069)	(0.052)	(0.046)	(0.054)	(0.054)
rsecondary	-0.034	0.068	0.084	0.069	-0.078	-0.058	-0.069	-0.070
	(0.077)	(0.063)	(0.083)	(0.082)	(0.064)	(0.057)	(0.068)	(0.068)
rhighabove	-0.127	0.029	0.032	0.009	-0.174*	-0.113	-0.129	-0.130
	(0.109)	(0.089)	(0.118)	(0.118)	(0.103)	(0.090)	(0.108)	(0.108)
rmarried	-0.643***	-0.661***	-0.771***	-0.769***	-0.388**	-0.283*	-0.322**	-0.333**
	(0.237)	(0.215)	(0.257)	(0.256)	(0.152)	(0.146)	(0.163)	(0.163)
hchild	-0.213***	-0.185***	-0.232***	-0.228***	-0.049	-0.037	-0.045	-0.045
	(0.059)	(0.052)	(0.064)	(0.064)	(0.045)	(0.043)	(0.048)	(0.048)
hhhres	-0.016	-0.005	-0.005	-0.005	0.019	0.018	0.019	0.018
	(0.023)	(0.021)	(0.025)	(0.025)	(0.016)	(0.016)	(0.017)	(0.017)
lnhhadurbl	-0.015	-0.019	-0.021	-0.022	0.000	-0.008	-0.006	-0.006
	(0.016)	(0.015)	(0.017)	(0.017)	(0.012)	(0.011)	(0.013)	(0.013)
y2015	0.096*	0.027	0.071	0.070	0.115***	0.089***	0.114***	0.116***
	(0.051)	(0.045)	(0.055)	(0.055)	(0.037)	(0.035)	(0.039)	(0.039)
urban_nbs	0.687***	0.554***	0.754***	0.744***	0.580***	0.621***	0.741***	0.747***
	(0.150)	(0.119)	(0.170)	(0.170)	(0.103)	(0.089)	(0.113)	(0.114)
ξ_i	1.000		1.000	1.000	1.000		1.000	1.000
	.		.	.	.		.	.
_cons	-2.709	-1.333	-0.823	-0.745	-1.619	-1.276	-1.638	-1.589
	(2.373)	(2.084)	(2.648)	(2.608)	(1.586)	(1.517)	(1.740)	(1.744)
Provinces	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AgeGroups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{x}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\sigma_3^{2(O)}$	0.553***		0.570***	0.567***	0.309***		0.299***	0.304***
	(0.201)		(0.211)	(0.207)	(0.116)		(0.115)	(0.115)
ICC	0.356***		0.363***	0.362***	0.236***		0.230***	0.233***
ρ_{13^O}				0.163**				-0.030
				(0.081)				(0.066)
ρ_{23^O}				0.360***				0.184***
				(0.066)				(0.058)
Observation	7803	6824	6824	6834	10057	9092	9092	9106
Log Likelihood	-2704	-2339	-2332	-19019	-4563	-4050	-4045	-28641

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Sample of individuals aged 40 or above at the first wave, staying for wave 1 and 2 and not missing in the outcome of L^O_{it} in column 1 and 5, or not missing in the outcomes of M_{it}, P_{it} and L^O_{it} for the rest of models, not belonging to the group of formal retirees or formal workers reaching the compulsory retirement ages, are selected for the estimation.

Table 3.10: Comparison of Models for Agricultural Work

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	RE	Pooled	RE	Tri	RE	Pooled	RE	Tri
L^W_{it-1}	0.488***	0.894***	0.520***	0.521***	0.553**	0.873***	0.582**	0.865***
	(0.157)	(0.081)	(0.187)	(0.186)	(0.262)	(0.113)	(0.272)	(0.245)
L^W_{i1}	1.101***	0.607***	1.143***	1.141***	1.423***	0.876***	1.328***	0.915***
	(0.216)	(0.083)	(0.262)	(0.260)	(0.402)	(0.116)	(0.408)	(0.343)
M_{it-1}		0.004	-0.000	-0.050		0.023	0.019	0.026
		(0.061)	(0.072)	(0.077)		(0.063)	(0.072)	(0.066)
P_{it-1}		0.004	0.006	0.004		-0.001	-0.001	0.003
		(0.012)	(0.015)	(0.016)		(0.014)	(0.016)	(0.015)
rrural2	0.512***	0.375***	0.469***	0.467***	0.715***	0.616***	0.714***	0.640***
	(0.125)	(0.102)	(0.132)	(0.132)	(0.149)	(0.110)	(0.149)	(0.133)
rage	0.112	0.125	0.153*	0.152*	0.193**	0.185**	0.198**	0.194**
	(0.085)	(0.077)	(0.091)	(0.091)	(0.090)	(0.082)	(0.094)	(0.083)
rage_sq	-0.074	-0.084	-0.103	-0.103	-0.143**	-0.133**	-0.143**	-0.139**
	(0.063)	(0.057)	(0.068)	(0.068)	(0.069)	(0.063)	(0.072)	(0.064)
rprimary	-0.025	0.006	0.014	0.013	-0.079	-0.059	-0.063	-0.059
	(0.082)	(0.070)	(0.087)	(0.087)	(0.089)	(0.075)	(0.089)	(0.076)
rsecondary	-0.223***	-0.175**	-0.223**	-0.220**	-0.270***	-0.166*	-0.184*	-0.179**
	(0.084)	(0.069)	(0.091)	(0.091)	(0.104)	(0.085)	(0.102)	(0.087)
rhighabove	-0.230**	-0.121	-0.155	-0.156	-0.592***	-0.488***	-0.560***	-0.505***
	(0.106)	(0.088)	(0.112)	(0.112)	(0.177)	(0.138)	(0.175)	(0.149)
rmarried	0.915***	0.852***	0.910**	0.934***	0.469	0.405	0.468	0.469
	(0.318)	(0.303)	(0.358)	(0.357)	(0.343)	(0.313)	(0.360)	(0.314)
hchild	0.010	0.033	0.043	0.039	0.104	0.093	0.115	0.076
	(0.078)	(0.072)	(0.085)	(0.086)	(0.084)	(0.072)	(0.084)	(0.073)
hhhres	-0.046	-0.047*	-0.055*	-0.055*	-0.013	-0.029	-0.033	-0.022
	(0.030)	(0.028)	(0.032)	(0.032)	(0.031)	(0.028)	(0.031)	(0.028)
lnhhadurbl	-0.049**	-0.046*	-0.049*	-0.049*	0.013	0.011	0.010	0.015
	(0.024)	(0.023)	(0.027)	(0.027)	(0.028)	(0.026)	(0.029)	(0.026)
y2015	-0.213***	-0.190***	-0.240***	-0.241***	-0.200***	-0.178***	-0.192***	-0.171***
	(0.060)	(0.053)	(0.066)	(0.066)	(0.068)	(0.063)	(0.071)	(0.064)
urban_nbs	-1.186***	-0.824***	-1.059***	-1.057***	-1.139***	-0.915***	-1.106***	-0.961***
	(0.257)	(0.197)	(0.268)	(0.268)	(0.280)	(0.192)	(0.287)	(0.243)
ξ_i	1.000		1.000	1.000	1.000		1.000	1.000
	.		.	.	.		.	.
_cons	-3.783	-4.354*	-5.250*	-5.215*	-6.889**	-6.686***	-7.209**	-7.022***
	(2.699)	(2.449)	(2.896)	(2.892)	(2.759)	(2.485)	(2.848)	(2.528)
Provinces	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AgeGroups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{x}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\sigma_3^{2(W)}$	0.470**		0.438*	0.439*	0.393		0.307	0.052
	(0.201)		(0.230)	(0.229)	(0.305)		(0.288)	(0.207)
ICC	0.320**		0.305*	0.305*	0.282		0.235	0.049
ρ_{13^W}				0.228**				-0.084
				(0.105)				(0.251)
ρ_{23^W}				0.036				-0.142
				(0.084)				(0.236)
Observation	5584	4895	4895	6834	5634	5213	5213	9612
Log Likelihood	-1981	-1659	-1656	-18364	-1230	-1126	-1125	-26882

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Sample of individuals aged 40 or above at the first wave, staying for wave 1 and 2 and not missing in the outcome of L^W_{it} in column 1 and 5, or not missing in the outcomes of M_{it}, P_{it} and L^W_{it} for the rest of models, not belonging to the group of formal retirees or formal workers reaching the compulsory retirement ages, are selected for the estimation.

labour market performance, and health that continue into adulthood and in older ages. Rural-hukou holders in our sample tend to receive less education and are more likely to do agricultural work when they grow up.

Rural-hukou holders might live in urban areas due to migration during adulthood or due to urbanization. The negative effect of being born or raised in rural areas on the risk of old-age depression cannot be fully compensated for by individuals currently living in urban areas ($urban_n bs$), as the effect of current, urban residence is insignificant for either men or women. The reason might be that rural hukou-holders living in urban areas are not entitled to the more generous and higher-quality urban welfare systems. Living in urban areas also does not significantly affect physical health conditions for men or women, but significantly predicts the probability of not working and of taking non-agricultural work. Non-agricultural workers tend to receive more generous pensions and stop working earlier than agricultural workers do. Urban residents have a higher chance to work in the non-agricultural sectors than rural residents do.

Age ($rage$) does not significantly predict the risk of depression or more physical health problems for either men or women, after controlling for other socio-economic status variables and the cross-effects of employment and other aspects of health. The age-squared variable ($rage_s q$) shows that women in senior age are in fact less likely to be depressed (Table 3.7), and are more likely to be not working (Table 9). Age is also marginally significant in predicting agricultural workers against non-agricultural workers for women (Table 3.10). The reason is that agricultural workers are in low socio-economic status, receive low or no retirement pensions, and tend to work into older ages than non-agricultural workers do.

The survey-year dummy ($y2015$) suggest that there is a significant time trend of higher depression rate and more physical health problems for both men and women in our sample, with little gender difference. Women are significantly more likely to stop working over time, but men are not (Table 3.9). There is a time trend of transiting from agricultural sector to non-agricultural sector among the current workers in our sample, more prevalent among men than women. The reason might be that women are more likely to choose to exit the labour force entirely instead of making sectoral changes. Men doing agricultural jobs may opt for less labour-intensive, non-agricultural jobs when they get old and gradually reduce working hours.

Educational attainments significantly predict better mental and physical health conditions, especially for women. Receiving secondary school education ($rsecondary$) significantly reduces the risk of depression and the number of physical problems for both men and women, compared to those receiving less than primary school education (the reference group). The effect of receiving high school or above education ($rhighabove$) on depression is similar to the effect of secondary education for men, but almost doubles the size for women. Receiving high school or above education shows only a slightly larger positive effect on physical health conditions of men and women, compared to the effect of attaining secondary education. The fact that women overall report lower educational achievement than men do (Table 3.2) might contribute to the gender-specific health inequality. The overall more significant health-education gradient for women suggests that improving women's educational level can narrow the gender-specific health gap. Increasing high school enrolment is likely to improve old-age physical health substantially for men.

Education is also significant in predicting the different job sectors that people work in. Receiving secondary education significantly increases the chance of taking non-agricultural jobs, similarly for men and women. Receiving high school or above education doubles the chance only for women, indicating that women need more education than men do to work in the non-agricultural sectors.

Marriage (*rmarried*), including living with a partner, significantly lowers the risk of depression for both men and women, with the effect being almost 35% higher for women than it is for men (Table 3.7). The effect is the largest among all socio-economic status factors in determining the risk of depression, suggesting that the lack of emotional supports from partners and close family members is one of the biggest risk factors for old-age depression, especially for women. Marital status also significantly reduces the reported number of physical health problems for women but not for men (Table 3.8). The economically inactive are found to be more likely to live without a partner, and the chance more than doubles for men than for women. It is probably because the economically inactive are older than others, men older than women, and are more likely to be widows. Living alone without a partner, instead of age, might better explain the higher depression rate among the economically inactive group compared with others. The probability of living with a partner is not significantly different for agricultural and non-agricultural workers.

The number of children (*hchild*) is not significantly related to health and labour supply outcomes of either men or women. The number of household members (*hhhres*) has a marginally significant, positive effect on men's mental health, but increases the reported number of physical problems for women. The economically inactive women are also more likely to live in big households. The results indicate that women take the main burden of caring for household members and doing house chores. Household wealth, as measured by the log of the value of household durable assets (*lnhhadurbl*) is negatively related to the number of physical health problems for men or women.

In summary, the estimated signs of the socio-economic status variables are in line with our expectations and support the related literature, while also provide new, gender-specific evidence based on the particular context and sample of the paper. For example, using wave 1 of CHARLS data and studying the correlation between socio-economic status and depressive symptoms, Lei et al. (2014*a*) also find that depressive symptoms are higher for women and are significantly associated with low education attainment and per capita expenditures, worse childhood health, being a recent widow, the number of chronic health problems.

3.6 Robustness Check

3.6.1 Different Types of Labour Market Transitions

In this section, we further study the distribution of the relationship between depression, physical health, and labour supply status by investigating specific demographic groups. It can provide us with a more complete picture of the mechanisms through which health affects labour supply and vice versa.

Table 3.11 compares the estimates from the dynamic RE models using different definitions of labour supply status as the dependent variable. Columns 1 and 3 compare agricultural workers ($L_{it}^{FO} = 1$) with non-workers ($L_{it}^{FO} = 0$), and we study how much health conditions affect the risk of stopping agricultural work. Finally, Columns 2 and 4 compare non-agricultural workers ($L_{it}^{NFO} = 1$) with non-workers ($L_{it}^{NFO} = 0$), and we study the relationship between health and the probability of moving from non-agricultural jobs to non-working.

Starting with the lagged dependent variables, we can see that all three labour force status are state-dependent. Women are more stable than men and are less likely to re-enter the labour market after they stop working. Agricultural workers are more stable than non-agricultural workers.

Cross-effects show that depression in the previous period increases the probability of doing agricultural work instead of non-agricultural work, as shown in column 1. Depression does not significantly affect women's labour supply decisions. Physical health problems, instead increases the probability of exiting the labour force for both men and women working in either agricultural or non-agricultural sectors. The effect is stronger for non-agricultural workers, who tend to have better retirement support or old-age savings and can retire when they feel like, for example, in the case of physical health deterioration. The effect is weaker for agricultural workers, who suffer from insufficient old-age support and tend to work until they are physically incapable. Men tend to take more labour-intensive work than women do, even when they both work in the agricultural sector, and are thus affected more by the physical health problems.

3.6.2 Test for Attrition

Attrition bias can be a problem in our estimation if response rates are positively associated with health status. We use simple variable addition tests, proposed by Verbeek and Nijman (1992), to test for attrition bias. Following Contoyannis, Jones

Table 3.11: Trivaraite Models for Agricultural or Non-Agricultural Exits

	Men		Women	
	(1)	(2)	(3)	(4)
	L_{it}^{FO}	L_{it}^{NFO}	L_{it}^{FO}	L_{it}^{NFO}
L_{it-1}^{FO}	0.608***		0.673***	
	(0.116)		(0.073)	
L_{i1}^{FO}	1.232***		0.909***	
	(0.161)		(0.092)	
L_{it-1}^{NFO}		0.718***		0.934***
		(0.192)		(0.197)
L_{i1}^{NFO}		1.045***		1.077***
		(0.240)		(0.280)
M_{it-1}	0.217***	-0.053	0.057	0.116
	(0.064)	(0.107)	(0.041)	(0.112)
P_{it-1}	-0.083***	-0.089***	-0.050***	-0.139***
	(0.010)	(0.017)	(0.006)	(0.023)
rrural	0.288*	-0.236*	0.179**	-0.151
	(0.163)	(0.125)	(0.086)	(0.128)
rage	0.085*	-0.161***	0.108***	-0.059
	(0.044)	(0.061)	(0.027)	(0.050)
rage_sq	-0.107***	0.080*	-0.111***	0.009
	(0.033)	(0.047)	(0.022)	(0.042)
rprimary	0.014	0.066	0.032	-0.130
	(0.070)	(0.118)	(0.054)	(0.128)
rsecondary	-0.100	0.115	-0.073	0.155
	(0.088)	(0.131)	(0.068)	(0.154)
rhighabove	-0.127	0.199	0.068	0.088
	(0.141)	(0.147)	(0.124)	(0.204)
rmarried	0.937***	0.642	0.445***	0.287
	(0.202)	(0.508)	(0.112)	(0.320)
hchild	0.073*	-0.058	0.002	0.249**
	(0.041)	(0.100)	(0.031)	(0.101)
hhhres	-0.010	0.109**	-0.028*	-0.009
	(0.021)	(0.043)	(0.014)	(0.041)
lnhhadurbl	-0.012	0.041	-0.006	0.063*
	(0.016)	(0.031)	(0.011)	(0.036)
y2015	-0.103*	0.087	-0.129***	0.072
	(0.058)	(0.102)	(0.040)	(0.103)
urban_nbs	-1.073***	-0.318*	-1.192***	0.124
	(0.278)	(0.175)	(0.149)	(0.163)
ξ_i	1.000	1.000	1.000	1.000
	.	.	.	.
_cons	-2.661*	5.830***	-3.422***	0.398
	(1.471)	(2.026)	(0.859)	(1.621)
σ^2	0.687***	0.074	0.353***	0.351*
	(0.160)	(0.137)	(0.075)	(0.200)
$\rho_{13^{FO/NFO}}$	-0.119***	-0.025	-0.016	0.090
	(0.041)	(0.063)	(0.025)	(0.060)
$\rho_{23^{FO/NFO}}$	-0.119***	-0.035	-0.053***	0.052**
	(0.023)	(0.030)	(0.011)	(0.023)
Observation	9479	9478	12724	2994
Log Likelihood	-25986	-24127	-39562	-8942

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Sample of individuals aged 40 or above at the first wave, staying for wave 1 and 2 and not missing in the health outcomes and the respective labour supply dependent variables, not belonging to the group of formal retirees or formal workers reaching the compulsory retirement ages, are selected for the estimation.

and Rice (2004), we include one sample dummy in the main model: an indicator of whether the individual stays for all three waves (wave 1-3), in an attempt to test for any significant health-related attrition. The reference group includes individuals selected for only waves 1 and 2, either due to attrition, missing information, or formal retirement.

The estimates of the sample indicator ($w123sample$) show that men suffering from less physical health issues (column 2 in Table 3.12) are more likely to stay for all three waves. Although the magnitudes of state dependence and cross-effects change slightly after controlling for the attrition sample, the significance remain.

3.7 Conclusion

Focusing on the trivariate dynamic cross-effects model and using data of a nation-wide, longitudinal survey, we study the sequential causality between mental health, physical health, and labour market exits and entries of older, informal workers in China, after accounting for the persistence of the health and labour supply status. We also identify the roles of state dependence, observed socio-economic variables, and time-invariant individual unobserved effects in explaining the persistence of these outcomes. Our sample comprises mainly of older people from low socio-economic status households and working in the informal sector, either agricultural or non-agricultural, in China. They receive low-level or no retirement pensions or unemployment compensation, are covered by less generous health insurance, and tend to work into older age or until they are physically incapable, compared to formal employees in developed countries or in the urban areas of China. We find a bidirectional but asymmetric relationship between depression and physical frailty. Physical health deterioration increases the probability of exiting the labour force, while depressed men are more likely to stay in the labour force and take non-agricultural jobs than the non-depressed, mainly driven by financial pressure.

The paper provides new empirical evidence that adds a new dimension of mental health to the traditional studies looking into the relationships between physical or self-reported health and labour market transitions of older workers. By including mental health in studying the health effect on employment, we reduce the unexplained variation in health effects and enrich our understanding of the topics. While chronic diseases or functional limitations represent a permanent health shock, depression sets as a transitory shock given its marginally significant state dependence, especially for men. Only physical health deterioration, or permanent health shock, is found to have a significant, negative effect on older, informal workers' labour participation. By exploiting the dynamic relationship between depression and labour force participation, we can better understand the mechanisms behind individual labour supply adjustments to health shock.

Table 3.12: Trivariate Models Accounting for Attrition

	Men				Women			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	M_{it}	P_{it}	L^O_{it}	L^W_{it}	M_{it}	P_{it}	L^O_{it}	L^W_{it}
M_{it-1}	0.049	0.014	-0.187***	-0.050	0.115	0.034**	-0.066	0.024
	(0.096)	(0.019)	(0.064)	(0.077)	(0.078)	(0.014)	(0.042)	(0.075)
M_{i1}	0.763***				0.799***			
	(0.102)				(0.088)			
P_{it-1}	0.067***	0.017***	0.087***	0.005	0.049***	0.015***	0.057***	0.004
	(0.009)	(0.005)	(0.012)	(0.016)	(0.007)	(0.004)	(0.007)	(0.017)
P_{i1}		0.173***				0.135***		
		(0.007)				(0.005)		
L^O_{it-1}			0.619***				0.642***	
			(0.155)				(0.105)	
L^O_{i1}			1.026***				0.858***	
			(0.217)				(0.140)	
L^W_{it-1}				0.522***				0.587**
				(0.187)				(0.274)
L^W_{i1}				1.142***				1.319***
				(0.261)				(0.410)
L^F_{it-1}	-0.102	0.009			-0.041	-0.029*		
	(0.067)	(0.027)			(0.051)	(0.017)		
L^{NF}_{it-1}	-0.263***	0.003			0.051	-0.052*		
	(0.086)	(0.033)			(0.079)	(0.028)		
rrural2	0.315***	-0.031	0.008	0.465***	0.154*	-0.050*	-0.120	0.712***
	(0.102)	(0.040)	(0.118)	(0.132)	(0.086)	(0.030)	(0.081)	(0.149)
rage	0.072	0.040	-0.044	0.152*	0.066	0.039**	0.001	0.197**
	(0.062)	(0.026)	(0.082)	(0.091)	(0.049)	(0.017)	(0.055)	(0.094)
rage_sq	-0.065	-0.025	0.074	-0.103	-0.064*	-0.026**	0.032	-0.142**
	(0.044)	(0.019)	(0.059)	(0.068)	(0.035)	(0.012)	(0.040)	(0.072)
rprimary	0.042	0.009	0.038	0.014	-0.099*	-0.038*	-0.001	-0.062
	(0.057)	(0.027)	(0.068)	(0.087)	(0.058)	(0.022)	(0.054)	(0.088)
rsecondary	-0.115*	-0.050	0.078	-0.218**	-0.174**	-0.086***	-0.071	-0.182*
	(0.066)	(0.032)	(0.082)	(0.091)	(0.070)	(0.028)	(0.068)	(0.101)
rhighabove	-0.097	-0.130***	0.020	-0.154	-0.275**	-0.079	-0.146	-0.557***
	(0.097)	(0.047)	(0.118)	(0.112)	(0.118)	(0.050)	(0.108)	(0.175)
rmarried	-0.583***	-0.120	-0.763***	0.935***	-0.785***	-0.057	-0.334**	0.465
	(0.218)	(0.079)	(0.255)	(0.357)	(0.161)	(0.050)	(0.163)	(0.358)
hchild	0.034	-0.013	-0.229***	0.038	-0.055	-0.023*	-0.049	0.117
	(0.049)	(0.017)	(0.064)	(0.086)	(0.044)	(0.013)	(0.048)	(0.083)
hhhres	-0.019	0.009	-0.001	-0.055*	0.005	0.016***	0.021	-0.033
	(0.021)	(0.006)	(0.025)	(0.032)	(0.017)	(0.005)	(0.018)	(0.031)
lnhhadurbl	-0.006	-0.010*	-0.022	-0.048*	0.003	-0.004	-0.006	0.010
	(0.015)	(0.005)	(0.017)	(0.026)	(0.012)	(0.004)	(0.013)	(0.030)
y2015	0.132***	0.211***	0.124**	-0.238***	0.182***	0.218***	0.140***	-0.195***
	(0.043)	(0.014)	(0.055)	(0.066)	(0.036)	(0.011)	(0.039)	(0.071)
urban_nbs	-0.096	0.083	0.688***	-1.061***	-0.259**	-0.093**	0.733***	-1.108***
	(0.155)	(0.067)	(0.170)	(0.269)	(0.120)	(0.047)	(0.114)	(0.288)
w123sample	-0.070	-0.176***	-0.606***	-0.074	-0.036	-0.047	-0.324***	0.009
	(0.091)	(0.042)	(0.110)	(0.138)	(0.080)	(0.031)	(0.078)	(0.156)
_cons	-2.598	-0.832	-0.419	-5.131*	-1.799	-0.424	-1.524	-7.194**
	(1.973)	(0.840)	(2.582)	(2.896)	(1.547)	(0.543)	(1.757)	(2.846)
Provinces	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
AgeGroups	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
$\bar{x}_i$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
σ^2	0.490***	0.212***	0.551***	0.438*	0.635***	0.212***	0.308***	0.300
	(0.121)	(0.013)	(0.206)	(0.229)	(0.115)	(0.008)	(0.116)	(0.288)
ρ_{12}	0.181***	0.181***	0.181***	0.182***	0.180***	0.180***	0.180***	0.181***
	(0.020)	(0.020)	(0.020)	(0.020)	(0.015)	(0.015)	(0.015)	(0.015)
$\rho_{13^{O/W}}$	0.084**	0.084**	0.084**	0.106**	-0.013	-0.013	-0.013	-0.013
	(0.042)	(0.042)	(0.042)	(0.049)	(0.029)	(0.029)	(0.029)	(0.051)
$\rho_{23^{O/W}}$	0.117***	0.117***	0.117***	0.011	0.038***	0.038***	0.038***	-0.017
	(0.022)	(0.022)	(0.022)	(0.026)	(0.012)	(0.012)	(0.012)	(0.023)
Observation	6834	6834	6834	6834	9106	9106	9106	9106
Log Likelihood	-18992	-18992	-18992	-18354	-28630	-28630	-28630	-25726

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level.

Understanding how much poor mental and physical health affects individual decisions to exit or stay in the labour force, and the health consequences of old-age labour supply, has important policy implications. It points to the potential role that health interventions in periods of mental or physical illness can play on prolonging quality working life, and to the way in which labour market policies can impact on population health. The highly significant and positive bidirectional effects between mental and physical health, and the existence of unobserved effects that determine both aspects of health, implicate on the spill-over effects of health interventions that target on one aspect of health. Health policies targeting on not only physical health problems but also mental health, including depression and cognitive functioning, can be more effective than those simply focusing on physical illness. State dependence of the health and labour supply behaviour also suggests that health interventions and labour market policies can have long-lasting effects on individual welfare.

Specifically, we find significant state dependence for depression, physical health problems, and labour market status. Persistence of depression and non-working is larger for women than for men. We control for unobservables that significantly and positively predict both depression and worse physical health, and unobservables that predict worse physical health and risk of non-working for both men and women in the trivariate models. For men, we also find unobservables that significantly and positively relate to both depression, non-working, and agricultural works. Not accounting for the endogeneity of the cross-effects would bias the estimation of the interrelationships. The direction of the bias depends on the correlations between the unobservables that determine different outcomes.

Nonetheless, some cross-effects are still significant after we control for the common unobserved effects in the trivariate models. Physical health deterioration increases the probability of exiting the labour force. The effect is larger than the initial condition or state dependence, and is subject to sectoral and gender-specific differences. It emphasizes the role of health limitations in labour participation decisions, especially for men who are more likely to take labour-intensive jobs. Given the same level of physical deterioration, depressed men are more likely to be staying in the labour force than the non-depressed, mainly driven by workers taking up agricultural jobs in the subsequent period. The reason can be financial stress or loss of income due to being out of the labour force. We do not find evidence of agricultural workers transiting into less physical-demanding, non-agricultural jobs after physical health deterioration or depression. Labour supply status does not significantly affect physical health conditions in the subsequent period, though men who are non-working or taking agricultural jobs are more likely to be depressed in the following wave than those doing non-agricultural jobs.

Physical health deterioration predicts a higher risk of depression in the subsequent period, larger for men than for women. The effect is only second to the initial condition in predicting depression risk, and much more significant than socio-economic status variables. The initial condition has the strongest explanatory power on predicting physical health, followed by time effects. Rural-hukou and its associated disadvantaged early-life circumstances, low educational achievement, living without a partner are all found to significantly and negatively affect old-age mental and physical health for both men and women. There is also a time trend of worse mental and physical health, especially for women, higher probability of not working only for women, and a transition from agricultural to non-agricultural sector among the current workers, more for men than for women.

Appendix A

Social Pension and Labour Supply Responses: Evidence from New Rural Social Pension in China

A.1 Additional Tables

A.1.1 Distribution of labour supply outcomes by age and gender

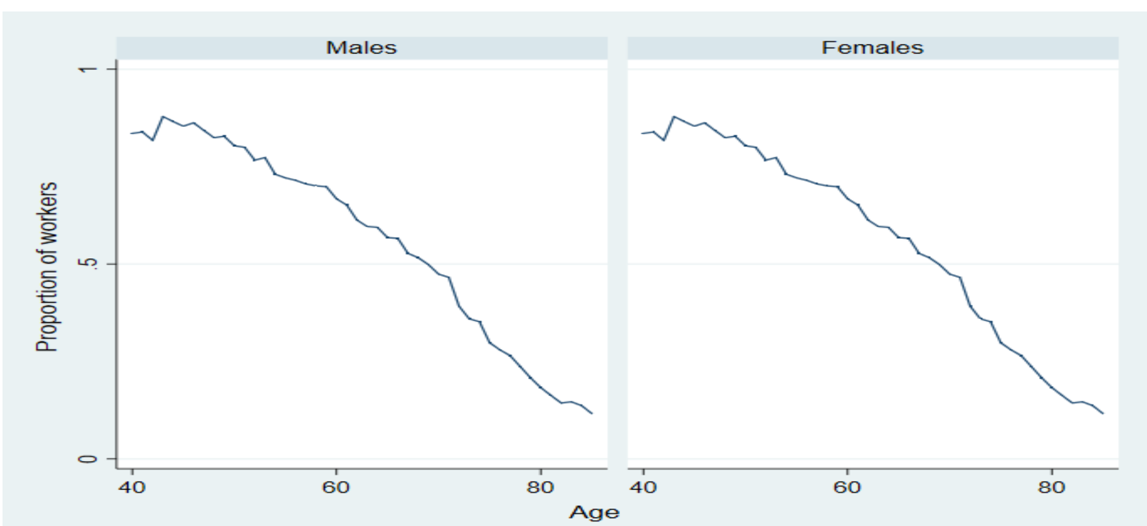

Figure A.1: Proportion of workers by age (restricted to sample aged 45-85)

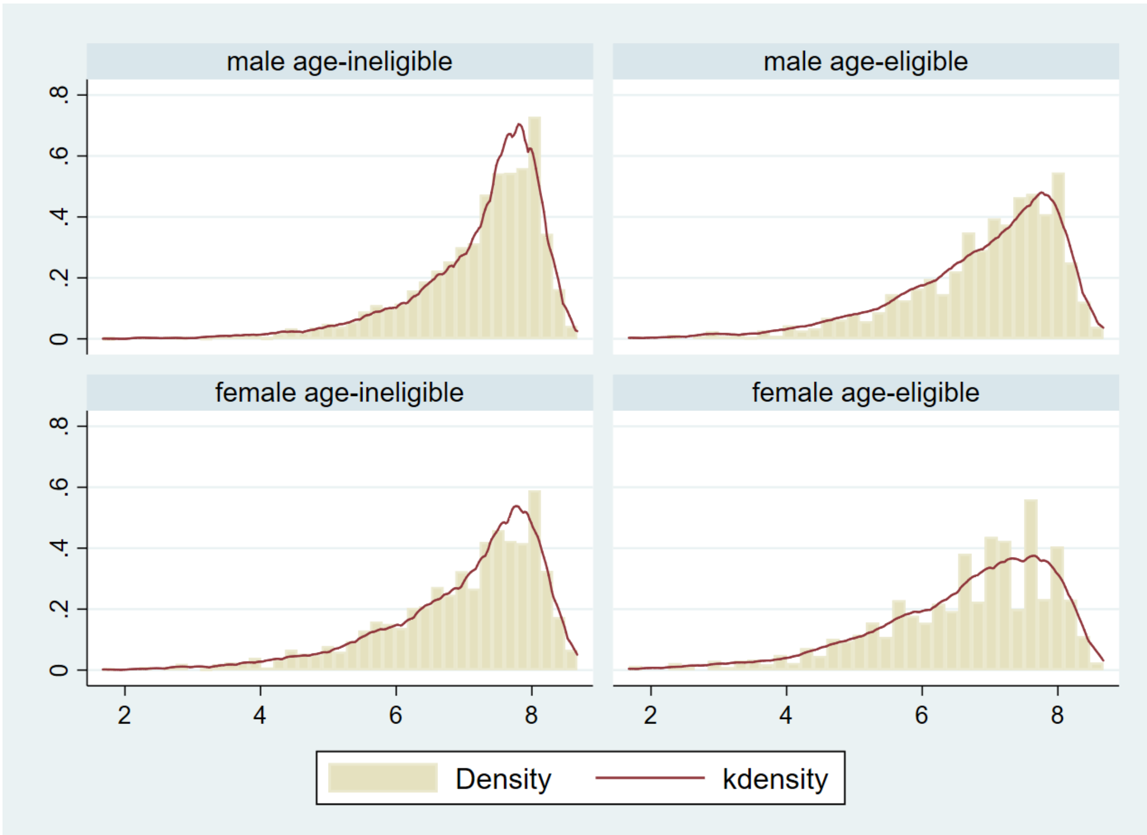

Figure A.2: Natural log of annual working hours by gender and age-eligibility

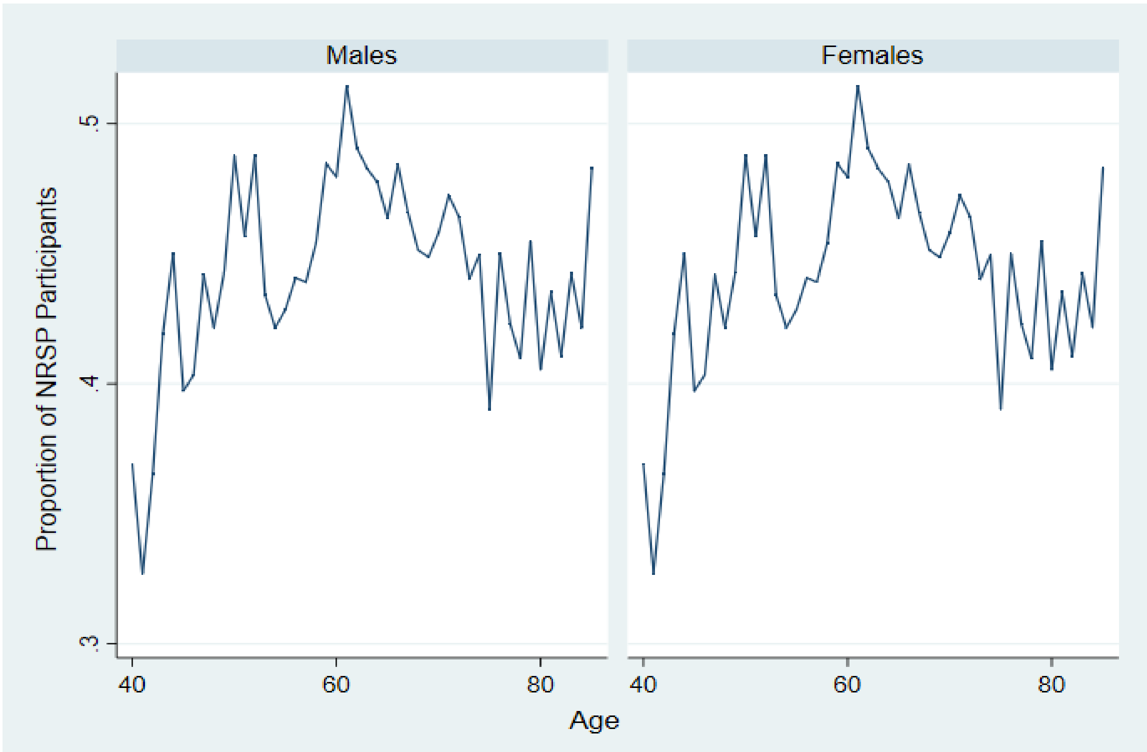

Figure A.3: Proportion of NRSP participants by age (restricted to sample aged 45-85)

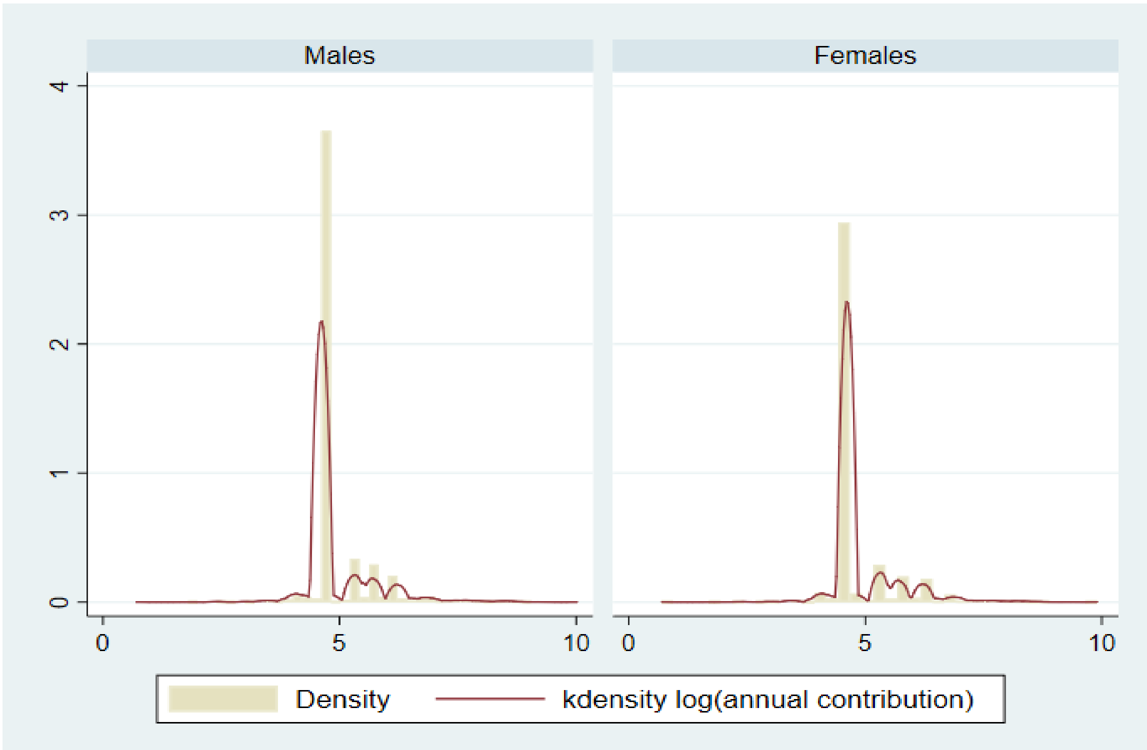

Figure A.4: Natural log of annual contributions of age-ineligible participants

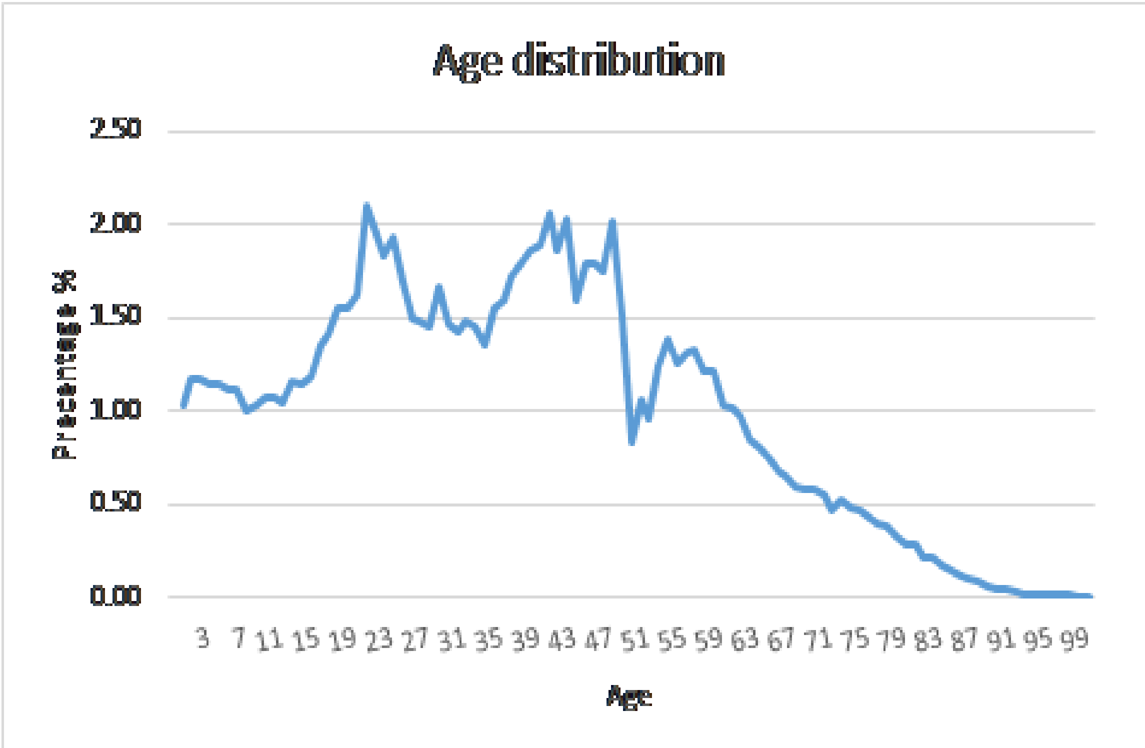

Figure A.5: Population age distribution

Table A.1: Determinants of NRSP Participation by Labour Supply Status

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	F-R	NF-Ret	F-NF	All	F-R	NF-Ret	F-NF	All
NRSP in place	2.050***	1.654***	1.935***	1.913***	2.016***	1.728***	1.948***	1.929***
	(0.151)	(0.077)	(0.058)	(0.053)	(0.053)	(0.073)	(0.057)	(0.048)
rage	0.102***	0.100***	0.104***	0.108***	0.092***	0.082***	0.100***	0.093***
	(0.015)	(0.016)	(0.016)	(0.012)	(0.011)	(0.013)	(0.014)	(0.010)
rage_sq	-0.078***	-0.073***	-0.076***	-0.080***	-0.068***	-0.057***	-0.073***	-0.068***
	(0.012)	(0.013)	(0.013)	(0.010)	(0.009)	(0.010)	(0.013)	(0.008)
rprimary	0.034	0.063	0.004	0.033	0.034	0.035	0.037	0.032
	(0.034)	(0.043)	(0.032)	(0.029)	(0.032)	(0.043)	(0.034)	(0.029)
rsecondary	0.064	0.055	0.043	0.057*	0.078*	0.036	0.093**	0.064*
	(0.043)	(0.047)	(0.036)	(0.033)	(0.043)	(0.054)	(0.042)	(0.038)
rhighabove	0.016	0.002	0.037	0.017	0.123	-0.057	0.074	0.057
	(0.072)	(0.063)	(0.050)	(0.047)	(0.076)	(0.086)	(0.074)	(0.063)
rmarried	0.180	-0.121	0.017	0.079	-0.062	-0.241*	-0.105	-0.125
	(0.128)	(0.161)	(0.147)	(0.117)	(0.101)	(0.133)	(0.124)	(0.096)
hchild	0.066**	0.079**	0.059**	0.070***	0.092***	0.065*	0.091***	0.084***
	(0.029)	(0.038)	(0.028)	(0.024)	(0.025)	(0.035)	(0.029)	(0.023)
hhhres	-0.013	-0.015	-0.018	-0.018	-0.000	-0.030*	0.004	-0.005
	(0.014)	(0.016)	(0.012)	(0.011)	(0.011)	(0.015)	(0.012)	(0.010)
lnhhadurbl	0.016	0.012	0.015	0.016**	0.012	0.003	0.023**	0.011
	(0.011)	(0.012)	(0.010)	(0.008)	(0.008)	(0.010)	(0.010)	(0.007)
y2013	0.473***	0.436***	0.444***	0.448***	0.452***	0.368***	0.436***	0.428***
	(0.069)	(0.044)	(0.032)	(0.029)	(0.030)	(0.043)	(0.032)	(0.027)
y2015	0.444***	0.379***	0.424***	0.414***	0.387***	0.292***	0.380***	0.365***
	(0.071)	(0.043)	(0.033)	(0.029)	(0.031)	(0.043)	(0.033)	(0.028)
urban_nbs	-0.053	-0.157*	-0.237**	-0.173**	-0.069	-0.083	-0.177*	-0.141**
	(0.172)	(0.086)	(0.094)	(0.081)	(0.088)	(0.074)	(0.099)	(0.067)
_cons	-5.284***	-5.379***	-5.176***	-5.466***	-4.805***	-4.351***	-5.086***	-4.823***
	(0.509)	(0.546)	(0.495)	(0.399)	(0.353)	(0.441)	(0.440)	(0.317)
$\rho_{NRSP-work}$	0.037	-0.064	0.132***	0.030	0.094**	-0.003	0.072	0.104***
	(0.058)	(0.082)	(0.051)	(0.051)	(0.039)	(0.070)	(0.055)	(0.038)
$\rho_{NRSP-hrs}$	-0.020	-0.120**	-0.068**	-0.068**	0.034	-0.061	0.016	0.006
	(0.056)	(0.058)	(0.031)	(0.032)	(0.035)	(0.068)	(0.029)	(0.028)
Observations	12825	8633	15385	18635	17675	9009	15070	21227
Log likelihood	-56381	-32571	-84953	-85569	-71793	-23453	-81999	-90001

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, belonging to the specific labour groups studied in each model, and not missing in the outcome variables.

Table A.2: Determinants of NRSP Participation by Educational Groups

	Males			Females		
	(1)	(2)	(3)	(4)	(5)	(6)
	illiterate	primary	secondary	illiterate	primary	secondary
NRSP in place	2.168***	1.883***	1.864***	1.860***	2.040***	1.893***
	(0.148)	(0.074)	(0.088)	(0.070)	(0.084)	(0.115)
rage	0.064**	0.114***	0.155***	0.080***	0.098***	0.130***
	(0.026)	(0.016)	(0.031)	(0.013)	(0.018)	(0.042)
rage_sq	-0.049**	-0.082***	-0.123***	-0.060***	-0.069***	-0.101***
	(0.020)	(0.013)	(0.026)	(0.011)	(0.016)	(0.038)
rmarried	-0.074	0.035	0.482**	-0.228*	0.058	-0.001
	(0.284)	(0.153)	(0.231)	(0.130)	(0.170)	(0.254)
hchild	0.029	0.069**	0.098**	0.122***	0.017	0.043
	(0.066)	(0.031)	(0.047)	(0.031)	(0.039)	(0.064)
hhhres	-0.020	-0.028*	-0.006	0.006	-0.014	-0.013
	(0.029)	(0.016)	(0.018)	(0.015)	(0.017)	(0.024)
lnhhadurbl	0.039**	0.001	0.032**	0.019*	0.002	0.000
	(0.019)	(0.012)	(0.015)	(0.010)	(0.013)	(0.021)
y2013	0.521***	0.516***	0.338***	0.555***	0.335***	0.318***
	(0.078)	(0.041)	(0.047)	(0.040)	(0.043)	(0.068)
y2015	0.361***	0.478***	0.354***	0.466***	0.314***	0.236***
	(0.082)	(0.042)	(0.051)	(0.043)	(0.044)	(0.072)
urban_nbs	0.202	-0.249**	-0.122	-0.142	-0.097	-0.295**
	(0.404)	(0.109)	(0.122)	(0.123)	(0.103)	(0.132)
_cons	-4.073***	-5.732***	-6.666***	-4.308***	-5.421***	-4.912***
	(1.108)	(0.537)	(0.934)	(0.476)	(0.573)	(1.174)
$\rho_{NRSP-work}$	0.131	-0.040	0.081	0.140**	0.091	-0.007
	(0.100)	(0.071)	(0.107)	(0.057)	(0.063)	(0.094)
$\rho_{NRSP-hrs}$	-0.139	-0.053	-0.101*	0.040	-0.017	-0.037
	(0.089)	(0.041)	(0.060)	(0.045)	(0.044)	(0.071)
Observations	2633	9652	6350	9548	8414	3261
Log likelihood	-10697	-44193	-30466	-37991	-36904	-14805

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, belonging to the specific educational groups studied in each model, and not missing in the outcome variables.

Table A.3: Determinants of NRSP Participation by Asset Quartiles

	Males				Females			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4
NRSP in place	1.761***	2.108***	1.928***	1.867***	1.866***	1.996***	2.041***	1.858***
	(0.096)	(0.112)	(0.106)	(0.112)	(0.092)	(0.097)	(0.101)	(0.101)
rage	0.094***	0.132***	0.113***	0.088***	0.114***	0.096***	0.090***	0.089***
	(0.023)	(0.026)	(0.025)	(0.028)	(0.020)	(0.019)	(0.019)	(0.021)
rage_sq	-0.068***	-0.102***	-0.085***	-0.063***	-0.083***	-0.072***	-0.064***	-0.065***
	(0.018)	(0.021)	(0.020)	(0.023)	(0.015)	(0.016)	(0.015)	(0.018)
rprimary	0.073	0.033	-0.011	0.041	-0.008	0.044	0.132**	-0.044
	(0.053)	(0.054)	(0.057)	(0.061)	(0.060)	(0.053)	(0.054)	(0.055)
rsecondary	0.091	0.001	0.064	0.068	0.144	0.074	0.146**	-0.023
	(0.066)	(0.061)	(0.061)	(0.064)	(0.094)	(0.071)	(0.067)	(0.061)
rhighabove	-0.127	0.032	0.081	0.028	0.215	0.033	0.199	-0.052
	(0.109)	(0.091)	(0.086)	(0.081)	(0.158)	(0.127)	(0.129)	(0.099)
rmarried	0.060	0.305	-0.099	-0.042	-0.290**	-0.078	0.220	-0.267
	(0.205)	(0.250)	(0.239)	(0.299)	(0.142)	(0.194)	(0.216)	(0.265)
hchild	0.031	0.065	0.121**	0.076	0.042	0.161***	0.084*	0.043
	(0.044)	(0.051)	(0.056)	(0.058)	(0.040)	(0.048)	(0.050)	(0.056)
hhhres	-0.020	-0.027	-0.001	-0.019	-0.000	-0.032	-0.004	0.022
	(0.023)	(0.026)	(0.024)	(0.022)	(0.021)	(0.023)	(0.021)	(0.020)
lnhhadurbl	0.034**	-0.027	-0.005	-0.029	0.014	0.072	-0.021	-0.009
	(0.013)	(0.068)	(0.085)	(0.042)	(0.012)	(0.062)	(0.082)	(0.039)
y2013	0.525***	0.587***	0.385***	0.364***	0.421***	0.469***	0.437***	0.402***
	(0.061)	(0.062)	(0.063)	(0.058)	(0.057)	(0.058)	(0.058)	(0.054)
y2015	0.427***	0.507***	0.458***	0.343***	0.347***	0.359***	0.379***	0.378***
	(0.062)	(0.066)	(0.067)	(0.061)	(0.059)	(0.062)	(0.063)	(0.056)
urban_nbs	0.116	-0.151	-0.270*	-0.206*	0.338**	-0.194	-0.215*	-0.250**
	(0.177)	(0.175)	(0.154)	(0.123)	(0.153)	(0.153)	(0.122)	(0.111)
_cons	-5.234***	-6.336***	-5.554***	-3.017***	-5.632***	-5.448***	-4.738***	-3.824***
	(0.764)	(0.925)	(1.038)	(0.920)	(0.662)	(0.727)	(0.887)	(0.763)
$\rho_{NRSP-work}$	0.042	0.130	0.077	-0.147	0.228***	-0.004	0.129*	0.014
	(0.101)	(0.110)	(0.121)	(0.110)	(0.080)	(0.076)	(0.076)	(0.077)
$\rho_{NRSP-hrs}$	-0.119	-0.047	-0.053	-0.070	0.060	-0.025	0.007	-0.004
	(0.089)	(0.060)	(0.055)	(0.065)	(0.063)	(0.058)	(0.056)	(0.058)
Observations	4664	4732	4666	4573	5312	5383	5262	5269
Log likelihood	-19968	-21893	-21797	-21593	-20454	-23038	-23248	-22918

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Standard errors are clustered in individual level. Other covariates not reported include the indicators of 7 types of administrative areas and of 28 provinces, individual-level means for the time-varying variables. The sample is restricted to rural residents not formally retired, belonging to the specific asset quantile groups studied in each model, and not missing in the outcome variables.

A.2 Fixed Effects Estimation

In this section, we re-estimate the trivariate models in the previous sections using alternative FE approaches to relieve the concern that the covariates might correlate with the unobserved individual heterogeneity. We estimate FE IV models separately for the working probability decision and the working hour decision. The community-level placement of the NRSP is used as an instrumental variable for the potentially endogenous variable of NRSP participation. To account for the sample selection bias when we study the working hours of the current workers, we use the panel heckit approach. Specifically, we estimate a random effect Probit model for the working probability, in which being currently working is a function of all covariates. We construct the inverse mill ratio and include it as an extra independent variable in the FE IV estimation of working hours. A significant inverse mill ratio has the same meaning as a significant $\rho_{work-hrs}$ in the trivariate model, and indicates that there are unobserved effects that predict both working probability and hours of working. Not accounting for the unobserved effects and estimating a linear model of working hours cause biased estimation of the covariates. We use the log of weekly working hours to study the percentage change.

Although the FE model controls for unobservable, time-invariant individual characteristics, it has several drawbacks. First of all, it does not account for the binary nature of the endogenous pension participation status. The two-stage FE IV model also does not account for the binary nature of the working status, and estimates a linear model. Secondly, although we include the inverse mill ratio, the model still does not account for cross-equation correlations of the random effect of the working probability model and that of the working hour model. Finally, multistage estimation is less efficient than maximum likelihood estimation and generates wrong standard errors. A bootstrapped standard error should be used.

A.2.1 Labour Supply Decision and NRSP Participation Decision

Tables A.4 and A.5 report FE estimation results for males and females, respectively. Models 1, 2 and 3 use a two-stage FE IV approach to estimate the ATE of NRSP participation on individual labour supply outcomes. Models 4 and 5 use a FE approach in which the possibly endogenous individual participation status is replaced with the exogenous, community-level NRSP status. In other words, models 4 and 5 estimate the ITT effect of NRSP-eligibility, or the average treatment effects across individuals actually enrolled and not enrolled in the NRPS. Specifically, model 4 is estimated by a linear FE model, while model 5 is estimated by a logistic FE model to account for the binary nature of the working status outcome. We compare their signs and the significant levels of ATE and ITT estimates in discussion.

labour supply responses may be more observable for rural workers staying in communities that started NRSP earlier, as it takes time for people to learn about and build commitment to the pension scheme. We estimate other FE ITT models (models 6 and 7) on community-level duration of the NRSP program, which is the difference between the year a community started the NRSP and a given survey year and captures both the availability of the NRSP and the time trend.

Table A.4: FE estimation for male participants

	FE IV			FE ITT			
	NRSP participants			NRSP placement		NRSP duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	hours	hours	work	hours	work	hours	work
Participants							
below 45	0.407	0.331	0.233*	0.160	0.208	0.070	0.011
	(0.330)	(0.277)	(0.135)	(0.132)	(0.470)	(0.067)	(0.136)
aged 45-49	0.061	0.085	0.077**	0.042	0.082*	0.031	-0.015
	(0.114)	(0.114)	(0.035)	(0.057)	(0.042)	(0.043)	(0.030)
aged 50-54	0.090	0.098	0.080**	0.061	0.095**	0.029	-0.015
	(0.100)	(0.100)	(0.032)	(0.050)	(0.037)	(0.039)	(0.025)
aged 55-59	0.042	0.043	0.038	0.031	0.053*	0.010	-0.026
	(0.095)	(0.095)	(0.031)	(0.047)	(0.030)	(0.040)	(0.025)
aged 60-64	-0.020	-0.027	0.022	-0.016	0.025	-0.005	-0.035
	(0.094)	(0.095)	(0.032)	(0.051)	(0.034)	(0.039)	(0.023)
aged 65-69	-0.033	-0.052	0.021	-0.037	0.026	-0.014	-0.034
	(0.109)	(0.107)	(0.035)	(0.060)	(0.037)	(0.040)	(0.022)
aged 70-74	-0.269**	-0.297**	-0.023	-0.220***	-0.004	-0.058	-0.043*
	(0.132)	(0.130)	(0.042)	(0.082)	(0.039)	(0.043)	(0.026)
aged 75-79	-0.387*	-0.515***	-0.131**	-0.355***	-0.090*	-0.103*	-0.072***
	(0.200)	(0.190)	(0.053)	(0.124)	(0.047)	(0.056)	(0.025)
Selection effect	-0.371						
	(0.258)						
Identification	522.13***	524.57***	739.05***				
Endogeneity	15.07*	18.23**	21.15**				
Observations	9547	9771	13865	11260	3887	11260	3887
N	3850	3962	5320	5423	1437	5423	3887
R^2 or LL	0.026	0.025	0.034	0.030	-1277.80	0.032	-1278.33

Notes: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Sample are restricted to rural workers not receiving urban employee pensions. Robust standard errors are used. We control for objective health measures of mobility score, indicators of heart disease, high blood pressure, dyslipidemia; spouses' work status, marrital status, household size, number of living parents, whether coresiding wit children, values of household durable goods, survey year dummies

Robust standard errors are used in all models. At the bottom of the table, we report the Lagrange Multiplier (LM) under identification test statistics, which confirms that community-level placement of the NRSP strongly predicts individual participation probability. We also report the endogeneity test statistics.

FE IV estimates tend to have similar signs and significant levels as the FE ITT estimates in both the working hours models (columns 2 and 4) and the working probability models (columns 3 and 5). Nonetheless, FE ITT estimates generally have smaller magnitudes than FE IV estimates, except for those produced by a logistic FE model (column 5). The logistic FE model selects on individuals whose working status changes over time, therefore has a small sample size and larger magnitude of estimates than linear probability models do.

First look at the male workers. As shown in Table A.9 above, endogeneity tests (models 1 to 3) reject the null hypothesis that the NRSP participation decision is exogenous to the labour supply decisions of male workers. The inverse mill ratio is not significant in model 1, suggesting that the working decision does not significantly relate to working hour decisions of male workers. The conclusion is consistent with the fact that the correlation of error terms is insignificant in the trivariate model (Table 1.6).

Both FE IV estimation and FE ITT estimation predict that participating in the pension program reduces working probability by more than 10% for male pensioners aged above 75, and by more than 15% for those aged above 80. The behaviours are likely to be driven by a pure income effect, as participants aged above 60 are eligible to receive basic pensions once they have been enrolled in the scheme. The expectation of an increase in future SSW, or future pension income, is also likely to affect current savings and labour supply. According to our theoretical model, the pension income substitutes for labour income in financing consumption and discourages labour supply after the pension-eligible age. Shu (2018) takes a RE IV approach on two waves of CHARLS data, and finds that receiving pension benefits only significantly increases the likelihood of stoping working for females by 27% , a magnitude much larger than those estimated here.

On the other hand, participants aged between 45 and 55 are about 8% more likely to enter the labour force after joining the NRSP scheme. Although benefiting from an increase in future SSW, young participants far away from the pension-eligible age would expect themselves to pay for pensions for the years to come. They may either take jobs to afford the minimum contributions, or to work more and save more on their pension accounts.

Column 7 suggests that male workers aged above 70 in communities that started the NRSP earlier are less likely to work and reduce working hours than their counterparts living in communities that started the NRSP later. Such a pattern is not observed among age-ineligible males. Nonetheless, the finding confirms the discouraging influence of receiving pensions on working probability and hours of working.

Estimations of the working hours models (models 2 and 4) suggest that for those already in the labour force, receiving a basic pension causes more than 20% a decline in hours of working for those aged above 70, and above 50% a drop in labour supply for those aged above 75. A life-cycle model can explain the larger responses of elder workers. Age-ineligible workers do not seem to change their working hours after participating in the pension program. The responses of age-eligible participants are likely to be driven by a pure income effect. Li, Wang and Zhao (2018) use two waves of CHARLS data and estimate a FE ITT model. They also find that age-eligible men reduce farm-working hours by 13.5%, a magnitude that is not far from our estimation.

Again, regression on the community level duration of the NRSP finds that the negative labour supply responses are larger for older participants in a community that starts the NRSP earlier.

Table A.5: FE estimation for female participants

	FE IV			FE ITT			
	NRSP participants			NRSP placement		NRSP duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	hours	hours	work	hours	work	hours	work
Participants							
below 45	0.221	0.166	-0.041	0.070	-0.033	-0.066	-0.031
	(0.206)	(0.201)	(0.069)	(0.098)	(0.071)	(0.067)	(0.046)
aged 45-49	0.010	-0.012	-0.032	-0.038	-0.035	-0.036	-0.007
	(0.120)	(0.119)	(0.038)	(0.058)	(0.040)	(0.048)	(0.029)
aged 50-54	0.107	0.118	-0.007	0.051	-0.005	-0.019	0.007
	(0.115)	(0.114)	(0.036)	(0.057)	(0.030)	(0.045)	(0.026)
aged 55-59	0.159	0.164	-0.027	0.082	-0.023	-0.037	0.008
	(0.107)	(0.108)	(0.035)	(0.055)	(0.030)	(0.046)	(0.025)
aged 60-64	0.139	0.126	-0.008	0.055	0.005	-0.059	0.014
	(0.111)	(0.111)	(0.034)	(0.060)	(0.029)	(0.045)	(0.026)
aged 65-69	0.129	0.119	0.018	0.053	0.034	-0.070	0.020
	(0.130)	(0.129)	(0.039)	(0.075)	(0.032)	(0.047)	(0.025)
aged 70-74	0.095	0.206	-0.064	0.114	-0.049	-0.048	-0.003
	(0.184)	(0.163)	(0.043)	(0.104)	(0.039)	(0.050)	(0.027)
aged 75-79	0.141	0.276	0.002	0.162	0.009	0.008	0.018
	(0.274)	(0.260)	(0.049)	(0.171)	(0.048)	(0.061)	(0.027)
Selection effect	0.469						
	(0.361)						
Identification	466.77***	467.12***	788.68***				
Endogeneity	7.97	8.72	3.87				
Observations	8729	8917	16200	10843	6030	10843	6030
N	3616	3710	6250	5593	2232	5593	6030
R^2 or LL	0.034	0.033	0.023	0.039	-2082.85	0.040	-2082.73

Notes: ibis.

Table A.5 displays the estimation results for females. Inverse mill ratio in model 1 is not significant, suggesting that there is no correlation between the working decision and the working hour decision. The conclusion is consistent with that of the error-correlation models in Table 1.7. Endogeneity tests (models 1 to 3) reject the null hypothesis that NRSP participation decision is exogenous to labour supply decisions of female workers.

Neither the FE IV estimation nor the FE ITT estimation finds any significant change in female workers' working probability or hours of working after they have participated into the pension program. Females living in communities starting the NRSP in different years also do not appear to behave in a systemically different way in their labour outcomes (models 6 and 7). Separating the estimation for different socio-demographic groups, for example, for workers in varying income quartiles or job sectors, or for workers having different commitment (contributing at different levels) to their pension account, may reveal some significant effects of the pension scheme on the behaviours of female workers.

A.2.2 Labour Supply Decision and NRSP Contribution Decision

In this section, we look at the group of age-ineligible participants who need to contribute to their pensions under the expectation of a higher future SSW. The aim is to study how the NRSP, apart from creating a pure income effect, also changes saving behaviours of participants and their labour market transition behaviours.

Specifically, we replace the key variable of interest of NRSP participation status with a continuous variable of natural log of annual contribution levels, and then redo the FE IV estimation using the subsample of NRSP participants aged below 60. To account for the possible endogeneity of the contribution decision, and to infer on a causal relationship between changes in contribution levels and labour supply responses, we use as instruments the community-level NRSP duration. The assumption is that the longer has the NRSP been in place, the more likely participants would build up commitment to it and save more on their pension accounts.

As mentioned earlier on, the distribution of the log of annual contributions is not normal, and about 75% of participants contribute at the minimum level of 100 RMB per year to their pension accounts. In models 4 and 5, we replace the continuous variable of contribution levels with an indicator of high contributor, which equals 1 for participants who contribute at a level higher than the minimum requirement of 100 RMB per year, and equals 0 for participants who contribute 100 RMB per year. More of our interest is the group of contributors who shift from contributing at the minimum requirement level to saving more for their pensions. These high contributors are likely to be systemically different from those contributing at the minimum level to the pension scheme in terms of commitment to the pension program and willingness to work.

Again, we report the estimation results separately for male (Table A.6) and female participants (Table A.7). As shown in Table A.6, the instruments are strong and significant based on under-identification test statistics at the bottom of the tables, and first stage estimation (not reported here). The inverse mill ratio is insignificant in model 1, suggesting that selecting contributors who are in the labour force to estimate changes in working hours does not significantly bias the estimation for male workers. Endogeneity tests fail to reject the null hypothesis that the contribution level decision is exogenous to the labour supply decision. Neither models looking at contribution levels nor those focusing on high contributors finds any significant causal relationship between contribution levels and labour supply responses.

Table A.6: FE estimation of contribution levels of male participants

	Contribution			High contributor	
	(1)	(2)	(3)	(4)	(5)
	hours	hours	work	hours	work
contributions of					
age below 45	-0.111	-0.171	-0.124	-0.382	-0.202
	(0.679)	(0.697)	(0.430)	(0.474)	(0.291)
age 45-49	-0.101	-0.149	-0.108	-0.236	-0.039
	(0.736)	(0.739)	(0.441)	(0.774)	(0.412)
age 50-54	-0.088	-0.147	-0.116	-0.237	-0.175
	(0.734)	(0.736)	(0.443)	(0.721)	(0.430)
age 55-59	-0.035	-0.109	-0.118	0.158	-0.157
	(0.726)	(0.728)	(0.443)	(0.635)	(0.412)
Selection effect	-1.999				
	(1.677)				
Identification	5.98**	5.95**	2.86*	11.41***	6.10**
Endogeneity	2.56	2.62	2.09	3.13	3.36
Observations	1883	1981	2431	1981	2431
N	867	916	1106	916	1106
R^2	0.020	0.016	-0.004	0.006	-0.019

Notes:

1. Sample are restricted to age-ineligible male participants

2. We control for objective health measures, spouses' working status, marrital status, household size, number of living parents, whether coresiding with children, values of household durable goods, survey year dummies.

3. Robust standard errors are used. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7 reports the labour supply estimation for female participants. The inverse mill ratio in model 1 is again insignificant. Endogeneity test statistics suggest that contribution levels are endogenous to the working hours decision of females but exogenous to the working decision. Nonetheless, the under-identification test fails to reject the null hypothesis that the instruments are too weak to identify the contribution levels of female participants for all models except for model 5. Therefore, we shall be cautious to interpret the estimation results.

Similar to the null patterns among male contributors, neither models looking at contribution levels nor those focusing on high contributors finds any significant change in labour supply outcomes following a change in contribution levels of female contributors.

Table A.7: FE estimation of contribution levels of female participants

	Contribution			High contributor	
	(1)	(2)	(3)	(4)	(5)
	hours	hours	work	hours	work
contributions of					
age below 45	2.056	2.416	-0.972	-1.064	-0.367
	(11.115)	(11.874)	(2.102)	(1.656)	(0.467)
age 45-49	2.100	2.466	-0.973	-1.089	-0.485
	(11.178)	(11.925)	(2.118)	(3.016)	(0.682)
age 50-54	2.166	2.536	-0.977	-0.846	-0.426
	(11.356)	(12.094)	(2.145)	(3.442)	(0.739)
age 55-59	2.177	2.564	-1.002	-1.345	-0.625
	(11.557)	(12.290)	(2.186)	(3.803)	(0.860)
Selection effect	1.322				
	(2.496)				
Identification	0.11	0.11	0.37	1.24	3.75*
Endogeneity	8.49*	7.88*	1.05	1.64	1.21
Observations	1879	1945	2894	1945	2894
N	874	907	1320	907	1320
R^2	-0.700	-0.959	-1.053	-0.050	-0.127

Notes: ibis.

According to our theoretical model, the expected future SSW is higher for high contributors, and should have had a larger offset effect on the labour income of these high contributors and of low income workers such as females and the elders. The FE IV estimations do not reveal such a pattern and changes in contribution levels are not found to be followed by any significant labour supply responses.

One explanation is that although the higher expected SSW provides a negative incentive for current labour supply, it is offset by a positive working incentive created by current contribution payments, and saving for economic inactivity. As shown in the participation models, male workers aged below 60 are found to retake jobs after joining the pension scheme. Female workers are not found to response significantly to the pension scheme, which should have benefited them more due to their relatively low labour income and higher dependence on private transfers for old-age support. Look into socio-demographic subgroups may reveal some patterns of significant labour supply responses among females. Finally, in line with the expectation of the theoretical framework, labour supply responses do have larger magnitudes among elder participants, who have lower uncertainty about future and survival risk, and smaller discount factor. The pension income, although still staying at low levels, exerts a larger substitution effect on the labour income of the elder workers.

To conclude, the NRSP that offers basic old-age security to the uncovered rural workers does to some extent improve the welfare of the target group. Specifically, as a substitute to labour income, pension incomes enable age-eligible workers to reduce workload or completely stop working earlier than they otherwise would do. As most of the informal workers take farming or temporary manufacturing jobs that are physically demanding and offer no security, they enjoy higher welfare in the sense that they are more confident of not working when they would like to. Although not so observable in the short-run, the NRSP also benefits younger workers as it introduces a tool of old-age savings that offers a much higher yield than deposit accounts, and it raises the awareness of saving and preparing for labour market inactivity. Financial literacy has been found to play an independent role in encouraging preparation for economic inactivity, which can be an increasing challenge in modern rural China where dependence on adult children for old-age support can hardly sustain due to the shrinking family size and urban migration.

A.2.3 Estimating Using the Rural Communities

In this section, we redo the FE estimation only on the subsample of workers in rural communities, and check whether the conclusion of the main discussions change under different assumptions of the NRSP placement.

Models 6 and 7 in Tables A.8 and A.9 suggest that workers located in rural communities that started the NRSP earlier do not response more in working probability or hours of working, in contrast to the estimation using the whole sample. Since rural communities generally started the NRSP earlier than urban communities did, the larger responses observed in communities having longer duration of the NRSP are likely driven by the systemically different responses of the NRSP participants located in rural and urban communities.

Compared to the estimation using the whole sample, the magnitudes tend to be larger for the significant estimates, but there are fewer significant estimates. Most importantly, the inverse mill ratio appears to be significantly negative, suggesting that the working hour decision is not independent of the working decisions for workers located in rural communities, and that neglecting the sample selection problem could result in a upward bias in the estimation. Therefore, we include the inverse mill ratio in the estimation of the ITT effects on working hours (Models 4 and 6). Participants at the age range of 70 to 75 are found to significantly reduce their hours of working by more than 20% (Models 2 and 4). The effect is similar to the estimation of the trivariate model.

Table A.8: FE estimation for male participants in rural communities

	FE IV			FE ITT			
	NRSP participants			NRSP placement		NRSP duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	hours	hours	work	hours	work	hours	work
Participants							
below 45	0.452	0.308	0.197	0.233	0.165	0.077	0.030
	(0.398)	(0.330)	(0.155)	(0.187)	(0.420)	(0.098)	(0.101)
aged 45-49	0.044	0.078	0.043	0.021	0.045	0.004	0.008
	(0.130)	(0.130)	(0.040)	(0.067)	(0.059)	(0.050)	(0.035)
aged 50-54	0.077	0.088	0.038	0.056	0.035	0.026	0.000
	(0.108)	(0.108)	(0.034)	(0.057)	(0.047)	(0.044)	(0.032)
aged 55-59	0.039	0.047	0.020	0.032	0.023	0.011	-0.005
	(0.103)	(0.103)	(0.034)	(0.052)	(0.041)	(0.045)	(0.029)
aged 60-64	-0.038	-0.044	0.029	-0.022	0.034	-0.011	-0.006
	(0.102)	(0.103)	(0.036)	(0.056)	(0.037)	(0.044)	(0.031)
aged 65-69	-0.000	-0.024	0.036	-0.022	0.039	-0.000	-0.008
	(0.118)	(0.117)	(0.040)	(0.066)	(0.038)	(0.045)	(0.031)
aged 70-74	-0.266*	-0.306**	-0.037	-0.192**	-0.030	-0.057	-0.024
	(0.152)	(0.151)	(0.048)	(0.094)	(0.042)	(0.048)	(0.032)
aged 75-79	-0.343	-0.544**	-0.240*	-0.381***	-0.116**	-0.097	-0.048
	(0.215)	(0.212)	(0.058)	(0.140)	(0.058)	(0.067)	(0.031)
Selection effect	-0.646**			-0.630**		-0.469	
	(0.300)			(0.299)		(0.316)	
Identification	430.74***	432.45***	586.17***				
Endogeneity	13.10	15.22*	16.64*				
Observations	7591	7755	10509	8529	2939	8529	2939
N	3054	3136	4049	3972	1092	3972	1092
R^2 or LL	0.026	0.023	0.030	0.029	-963.77	0.032	-963.73

Notes:

1. Sample are restricted to rural workers locating in rural communities, and not receiving urban employee pensions.
2. We control for objective health measures of mobility score, indicators of heart disease, high blood pressure, dyslipidemia; spouses' work status, marital status, household size, number of living parents, whether coresiding with children, values of household durable goods, survey year dummies
3. Robust standard errors are used. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9: FE estimation for female participants in rural communities

	FE IV			FE ITT			
	NRSP participants			NRSP placement		NRSP duration	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	hours	hours	work	hours	work	hours	work
Participants							
below 45	0.194	0.149	-0.045	0.061	-0.036	-0.100	0.007
	(0.209)	(0.206)	(0.071)	(0.108)	(0.116)	(0.073)	(0.058)
aged 45-49	0.011	-0.012	-0.042	-0.042	-0.036	-0.056	0.020
	(0.133)	(0.132)	(0.043)	(0.065)	(0.043)	(0.052)	(0.031)
aged 50-54	0.158	0.159	-0.015	0.079	-0.001	-0.018	0.025
	(0.126)	(0.125)	(0.041)	(0.064)	(0.032)	(0.049)	(0.026)
aged 55-59	0.135	0.145	-0.077**	0.072	-0.067*	-0.043	0.010
	(0.115)	(0.116)	(0.039)	(0.060)	(0.037)	(0.051)	(0.029)
aged 60-64	0.134	0.105	-0.012	0.040	0.014	-0.061	0.034
	(0.121)	(0.120)	(0.037)	(0.066)	(0.041)	(0.050)	(0.025)
aged 65-69	0.158	0.154	-0.012	0.077	0.019	-0.065	0.029
	(0.146)	(0.146)	(0.044)	(0.083)	(0.039)	(0.051)	(0.027)
aged 70-74	0.082	0.217	-0.093*	0.124	-0.066	-0.045	0.005
	(0.202)	(0.180)	(0.049)	(0.116)	(0.044)	(0.056)	(0.027)
aged 75-79	-0.151	-0.006	-0.017	-0.035	0.000	-0.052	0.033
	(0.303)	(0.289)	(0.057)	(0.182)	(0.061)	(0.064)	(0.031)
Selection effect	0.570						
	(0.409)						
Identification	388.09***	387.89***	625.84***				
Endogeneity	8.88	10.24	6.26				
Observations	6923	7071	12114	8486	4567	8486	4567
N	2854	2928	4687	4307	1704	4307	1704
R^2 or LL	0.034	0.033	0.025	0.039	-1560.92	0.040	-1561.61

Notes: ibis.

Recall that there is a null pattern of any pension effect in the FE estimations for the whole sample of female workers. Table A.9 reports the estimations that limit the sample to female workers in rural communities. Female participants at the age of 55 to 60 appear to be about 7% significantly less likely to work after joining the pension scheme.

In summary, FE estimations using only rural communities do not deviate too much from the whole sample FE estimations in terms of labour supply responses of elder participants aged above 70, but do have different conclusions in terms of behaviours of the age-ineligible participants and the timing of responses. Specifically, estimations using a restricted sample do not find any significant responses from the age-ineligible males, but do find females approaching pension-eligible age less likely to work. Moreover, estimations using a restricted sample fail to detect any significant timing of labour supply responses among male workers. Limiting the sample to rural communities produces a FE estimation closer to the estimations of trivariate models that control for community characteristics and individual means of observables (Models 4 and 8 in Tables 1.8 and 1.9).

This indicates that we should control for community-level characteristics in our model. We also try clustering standard errors on the community level instead of separating the estimation for rural communities, but the results (not reported here) are not very much different from the whole sample estimation using robust standard errors in Tables 9 and 10. A re-estimation of the FE IV model on contributors in rural communities finds no pattern of any significant change in labour outcomes after male or female contributors adjusting their level of savings on pension accounts.

Appendix B

Retirement Effect on Cognitive Functioning and Depression Risk of Formal Workers

B.1 AB Estimation Results

B.1.1 Effect of Transition into Retirement

The lagged dependent variables control for potential state-dependence of the health outcomes and are not significant in all cases. It can be due to finite sample bias, or that Y_{i1} is only weakly correlated with ΔY_{i2}.

As shown in the AB IV estimation in Table B.2, compared with the RE IV estimation in Table 2.3, we can see that the signs of retirement are consistent for men, but the retirement effect is only significant and positive in predicting men's episodic memory (column 3). The sizes of the effects estimated by the AB IV models are substantially larger than those by the bivariate RE estimation. The AB IV estimations of the retirement effect are insignificant across all three cognitive test outcomes for women. Apart from not accounting for state-dependence, the bivariate RE estimations also differ from the AB IV estimations in the estimation sample. Sample size is larger for the RE models than it is for the AB IV models, which select on individuals staying for all three waves.

Compared with AB models not accounting for endogeneity of the retirement status in Table B.3, state-dependence as represented by Y_{it-1} declines in significant levels and magnitudes, similarly for both men and women. The decline in the effect of Y_{it-1} could indicate that formal retirees who retire after retirement ages are more likely to stay in all three waves than other retirees do. It can also be that formal retirees' scores in cognitive tests are more consistent or show stronger state-dependence than

Table B.1: Retirement Age

	Men		Women	
	Freq.	Percent	Freq.	Percent
40	2	0.11	22	1.10
41	10	0.53	27	1.35
42	17	0.90	16	0.80
43	7	0.37	20	1.00
44	8	0.42	29	1.45
45	17	0.90	119	5.97
46	24	1.27	83	4.16
47	31	1.64	86	4.31
48	56	2.96	89	4.46
49	27	1.43	110	5.52
50	54	2.86	704	35.31
51	42	2.22	141	7.07
52	50	2.65	96	4.81
53	49	2.59	34	1.71
54	55	2.91	39	1.96
55	219	11.59	185	9.28
56	116	6.14	62	3.11
57	70	3.70	31	1.55
58	57	3.02	32	1.60
59	104	5.50	18	0.90
60	669	35.40	18	0.90
61	111	5.87	5	0.25
62	36	1.90	4	0.20
63	16	0.85	4	0.20
64	12	0.63	8	0.40
65	18	0.95	4	0.20
66	3	0.16	2	0.10
67	5	0.26	3	0.15
68	3	0.16	2	0.10
70	2	0.11	1	0.05
Total	1,890	100	1,994	100

scores of other retirees.

Table B.2: AB IV Estimation of Retirement Effect on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	3.049	2.277	4.564**	2.041	2.730	-0.093
	(1.992)	(1.468)	(2.126)	(2.022)	(2.407)	(1.904)
L.rimrc	0.103	-0.063				
	(0.094)	(0.087)				
L.rdlrc			0.067	0.081		
			(0.082)	(0.095)		
L.rmentalintact					0.083	-0.045
					(0.104)	(0.093)
rage_sq	-0.323	-0.163	0.146	0.317	-0.701**	-0.299
	(0.258)	(0.269)	(0.292)	(0.347)	(0.294)	(0.331)
rmarried	-0.249	0.335	-0.267	0.500	0.246	0.235
	(0.396)	(0.495)	(0.403)	(0.516)	(0.717)	(0.587)
hchild	-0.005	-0.114	0.158	-0.109	-0.118	-0.108
	(0.189)	(0.149)	(0.216)	(0.169)	(0.269)	(0.149)
hhhres	-0.051	-0.044	0.042	-0.111	0.082	0.031
	(0.077)	(0.074)	(0.083)	(0.093)	(0.092)	(0.096)
lnhhadurbl	0.011	0.071*	-0.089**	0.133**	0.051	0.067
	(0.039)	(0.041)	(0.044)	(0.052)	(0.065)	(0.050)
y2013	-0.408	-0.305	1.121	1.197	-1.461*	-0.352
	(0.688)	(0.682)	(0.792)	(0.904)	(0.754)	(0.825)
_cons	16.373	9.458	-6.003	-11.678	35.700***	19.577
	(11.955)	(11.357)	(13.518)	(14.994)	(13.244)	(13.635)
N	453	411	457	409	472	421
N_g	453	411	457	409	472	421
chi2	19.264	9.249	35.035	17.921	13.860	16.456

Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as extra instruments for retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, and having stayed for all three waves.

As shown in Table B.4, state-dependence as measured by the lagged dependent variable Y_{it-1} is insignificant for either men's or women's mental health and physical health. It means that for them, being diagnosed with depression symptoms during the previous wave or the number of physical problems reported in the previous wave is not significantly related to the risk of depression or the number of physical problems in the current wave.

Table B.3: AB Estimation of Retirement Effect on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.789**	-0.128	0.321	-0.037	0.237	-1.255**
	(0.357)	(0.318)	(0.371)	(0.407)	(0.313)	(0.622)
L.rimrc	0.112	-0.073				
	(0.096)	(0.086)				
L.rdlrc			0.065	0.071		
			(0.081)	(0.094)		
L.rmentalintact					0.068	-0.036
					(0.098)	(0.091)
rage_sq	-0.382	-0.322	-0.051	0.160	-0.805**	-0.247
	(0.249)	(0.256)	(0.274)	(0.310)	(0.314)	(0.354)
rmarried	-0.203	0.366	-0.166	0.509	0.301	0.227
	(0.393)	(0.492)	(0.387)	(0.517)	(0.710)	(0.582)
hchild	-0.038	-0.122	0.113	-0.105	-0.165	-0.101
	(0.190)	(0.144)	(0.212)	(0.167)	(0.262)	(0.143)
hhhres	-0.072	-0.033	0.003	-0.094	0.062	0.037
	(0.071)	(0.072)	(0.078)	(0.092)	(0.088)	(0.094)
lnhhadurbl	0.022	0.069*	-0.066	0.131**	0.061	0.063
	(0.039)	(0.041)	(0.042)	(0.053)	(0.065)	(0.051)
y2013	-0.628	-0.759	0.451	0.760	-1.822**	-0.240
	(0.643)	(0.639)	(0.720)	(0.785)	(0.799)	(0.871)
_cons	20.562*	17.850*	5.626	-3.690	42.250***	18.434
	(11.085)	(10.428)	(12.184)	(12.601)	(13.822)	(14.068)
N	453	411	457	409	472	421
N_g	453	411	457	409	472	421
chi2	20.059	6.966	30.849	17.008	13.741	19.701

Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, and having stayed for all three waves.

The estimated retirement effects are not significant for all models. Coefficients of other covariates are also not very significant in predicting men or women's mental or physical health, except for the age effect.

AB estimations not accounting for the endogeneity of the retirement decision (Table B.5) show that state-dependence as represented by Y_{it-1} declines in significant levels and magnitudes, similarly for both men and women. It suggests that retirees who retire after the retirement ages report more consistently on mental and physical health status than the other retirees. It can also be the fact that they are more likely to stay for all three waves, out of the reasons that we explained before.

B.1.2 Effects of Years in Retirement

State-dependence shows similar magnitudes and significant levels as they are in the retirement models for both men and women. Male retirees staying for longer time in retirement scores better in memory test (column 3) than those staying for fewer years or non-retirees do.

Table B.4: AB IV Estimation of Retirement Effect on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	-7.227	0.694	-0.671	0.087	-1.607	-0.962
	(4.956)	(4.811)	(0.483)	(0.494)	(1.510)	(1.333)
L.rcesd10	-0.081	-0.134				
	(0.096)	(0.095)				
L.rdepress			0.130	0.039		
			(0.097)	(0.095)		
L.rphyhealth					-0.134	-0.151
					(0.333)	(0.241)
rage_sq	0.846*	-0.475	0.053	0.004	0.807***	0.237
	(0.497)	(0.691)	(0.048)	(0.060)	(0.232)	(0.340)
rmarried	-1.400	-1.788*	-0.121	0.012	-0.053	-0.086
	(1.211)	(1.066)	(0.090)	(0.031)	(0.837)	(0.385)
hchild	-0.855*	-0.025	0.007	0.006	0.005	-0.099
	(0.505)	(0.401)	(0.032)	(0.026)	(0.109)	(0.099)
hhhres	-0.105	0.010	-0.021	-0.006	-0.073	0.028
	(0.142)	(0.261)	(0.014)	(0.022)	(0.078)	(0.077)
lnhhadurbl	-0.035	0.091	-0.011	0.004	-0.021	0.038
	(0.099)	(0.121)	(0.010)	(0.009)	(0.046)	(0.047)
y2013	2.054	-1.155	0.125	0.013	1.301**	-0.143
	(1.360)	(1.731)	(0.126)	(0.154)	(0.510)	(0.734)
_cons	-22.692	26.358	-1.489	-0.047	-29.882***	-4.198
	(23.398)	(28.869)	(2.182)	(2.539)	(9.839)	(13.324)
N	481	436	481	436	552	477
N_g	481	436	481	436	552	477
chi2	16.061	6.564	9.321	0.564	84.893	52.754

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, and having stayed for all three waves.

Table B.5: AB Estimation of Retirement Effect on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retire	0.408	0.196	-0.032	0.248*	-0.063	0.267
	(0.874)	(1.396)	(0.075)	(0.141)	(0.264)	(0.336)
L.rcesd10	-0.082	-0.135				
	(0.095)	(0.097)				
L.rdepress			0.098	0.051		
			(0.091)	(0.094)		
L.rphyhealth					-0.198	-0.203
					(0.336)	(0.257)
rage_sq	1.123**	-0.507	0.086*	0.014	1.078***	0.422
	(0.511)	(0.665)	(0.051)	(0.061)	(0.292)	(0.360)
rmarried	-1.661	-1.763*	-0.136	0.014	-0.054	-0.106
	(1.200)	(1.067)	(0.089)	(0.031)	(0.805)	(0.385)
hchild	-0.741	-0.018	0.016	0.005	0.014	-0.100
	(0.481)	(0.402)	(0.030)	(0.026)	(0.104)	(0.099)
hhhres	-0.040	0.007	-0.017	-0.008	-0.066	0.010
	(0.141)	(0.258)	(0.013)	(0.022)	(0.073)	(0.081)
lnhhadurbl	-0.106	0.075	-0.013	0.003	-0.036	0.035
	(0.095)	(0.124)	(0.010)	(0.010)	(0.045)	(0.048)
y2013	3.088**	-1.273	0.234*	0.040	2.016***	0.328
	(1.337)	(1.632)	(0.133)	(0.154)	(0.666)	(0.767)
_cons	-40.021*	28.152	-3.397	-0.592	-42.634***	-12.399
	(22.841)	(26.608)	(2.266)	(2.449)	(12.417)	(13.849)
N	481	436	481	436	552	477
N_g	481	436	481	436	552	477
chi2	16.256	5.926	8.718	3.842	88.109	51.763

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors, and having stayed for all three waves.

Table B.6: AB IV Estimation of Years in Retirement on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.192	-0.108	0.418*	-0.495	0.078	-0.394
	(0.171)	(0.290)	(0.216)	(0.359)	(0.216)	(0.339)
L.rimrc	0.134	-0.110				
	(0.100)	(0.085)				
L.rdlrc			0.086	0.012		
			(0.094)	(0.098)		
L.rmentalintact					0.069	-0.047
					(0.099)	(0.096)
rage_sq	-0.707*	-0.245	-0.512	0.578	-0.937*	0.147
	(0.384)	(0.421)	(0.423)	(0.485)	(0.512)	(0.618)
rmarried	-0.217	0.351	-0.060	0.451	0.467	0.166
	(0.420)	(0.509)	(0.425)	(0.532)	(0.745)	(0.608)
hchild	-0.027	-0.100	0.080	-0.098	-0.214	-0.128
	(0.195)	(0.154)	(0.224)	(0.171)	(0.268)	(0.151)
hhhres	-0.064	-0.023	0.011	-0.128	0.066	0.051
	(0.076)	(0.076)	(0.082)	(0.099)	(0.093)	(0.098)
lnhhadurbl	0.002	0.086**	-0.080*	0.119**	0.059	0.070
	(0.041)	(0.041)	(0.046)	(0.053)	(0.067)	(0.051)
y2013	-1.178	-0.789	-0.151	0.918	-2.073*	0.007
	(0.822)	(0.727)	(0.909)	(0.839)	(1.074)	(1.083)
_cons	33.583**	15.929	22.311	-13.973	47.337**	6.577
	(15.810)	(14.438)	(17.249)	(16.288)	(20.825)	(21.231)
N	435	384	439	382	451	391
N_g	435	384	439	382	451	391
chi2	19.315	9.410	28.613	16.294	12.210	14.529

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The distance to the mandatory retirement ages in years, constructed as the time elapsed since an individual reached the normal retirement age, which is 55 or 60 ($Agediff55_{it}$ and $Agediff60_{it}$) for men and 50 or 55 ($Agediff50_{it}$ and $Agediff55_{it}$) for women, are used as extra instruments for years in retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors and having stayed for all three waves.

Table B.7: AB Estimation of Years in Retirement on Cognitive Functioning

	IMR		DWR		Intact	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.043	-0.024	-0.026	-0.059	0.024	-0.140*
	(0.058)	(0.033)	(0.046)	(0.051)	(0.044)	(0.083)
L.rimrc	0.120	-0.102				
	(0.100)	(0.087)				
L.rdlrc			0.064	0.050		
			(0.088)	(0.100)		
L.rmentalintact					0.085	-0.032
					(0.100)	(0.094)
rage_sq	-0.570**	-0.363	-0.035	0.223	-0.937***	-0.174
	(0.290)	(0.267)	(0.309)	(0.329)	(0.354)	(0.385)
rmarried	-0.199	0.368	-0.081	0.516	0.440	0.215
	(0.414)	(0.494)	(0.406)	(0.521)	(0.749)	(0.585)
hchild	-0.034	-0.131	0.062	-0.090	-0.199	-0.152
	(0.196)	(0.145)	(0.219)	(0.169)	(0.273)	(0.145)
hhhres	-0.067	-0.020	0.007	-0.126	0.063	0.045
	(0.073)	(0.076)	(0.080)	(0.098)	(0.091)	(0.098)
lnhhadurbl	0.012	0.080*	-0.058	0.123**	0.063	0.057
	(0.041)	(0.042)	(0.043)	(0.054)	(0.069)	(0.051)
y2013	-1.060	-0.932	0.424	0.829	-2.154**	-0.335
	(0.718)	(0.654)	(0.801)	(0.818)	(0.890)	(0.913)
_cons	28.890**	19.681*	5.393	-5.277	47.640***	16.305
	(12.629)	(10.743)	(13.568)	(13.166)	(15.388)	(14.994)
N	435	384	439	382	451	391
N_g	435	384	439	382	451	391
chi2	18.102	9.714	28.031	16.822	13.298	16.511

Notes: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors and having stayed for all three waves.

Table B.8: AB IV Estimation of Years in Retirement on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.051	0.858	-0.028	0.052	-0.210	-0.115
	(0.342)	(0.781)	(0.032)	(0.078)	(0.140)	(0.243)
L.rcesd10	-0.100	-0.089				
	(0.101)	(0.097)				
L.rdepress			0.063	0.076		
			(0.090)	(0.103)		
L.rphyhealth					-0.227	-0.142
					(0.258)	(0.224)
rage_sq	0.843	-1.383	0.102	-0.047	1.297***	0.476
	(0.658)	(1.051)	(0.067)	(0.089)	(0.410)	(0.501)
rmarried	-2.084*	-1.851*	-0.160*	0.019	-0.170	-0.163
	(1.206)	(1.077)	(0.092)	(0.033)	(0.863)	(0.389)
hchild	-0.685	0.143	0.005	0.014	0.050	-0.064
	(0.485)	(0.395)	(0.029)	(0.025)	(0.108)	(0.102)
hhhres	-0.022	0.014	-0.013	-0.008	-0.067	0.014
	(0.144)	(0.271)	(0.013)	(0.024)	(0.076)	(0.080)
lnhhadurbl	-0.130	0.036	-0.014	0.003	-0.035	0.015
	(0.099)	(0.126)	(0.010)	(0.010)	(0.046)	(0.043)
y2013	2.417*	-1.726	0.234	-0.018	2.274***	0.345
	(1.445)	(1.784)	(0.144)	(0.155)	(0.806)	(0.844)
_cons	-27.141	52.601	-3.787	1.403	-49.974***	-13.143
	(27.307)	(35.541)	(2.744)	(2.969)	(16.383)	(17.570)
N	460	404	460	404	526	438
N_g	460	404	460	404	526	438
chi2	15.386	6.714	7.096	1.698	82.893	39.885

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Mandatory retirement ages, $d(Age_{it} \geqslant 50)$ and $d(Age_{it} \geqslant 55)$ for women, or $d(Age_{it} \geqslant 55)$ and $d(Age_{it} \geqslant 60)$ for men, are used as instruments for retirement. The distance to the mandatory retirement ages in years, constructed as the time elapsed since an individual reached the normal retirement age, which is 55 or 60 ($Agediff55_{it}$ and $Agediff60_{it}$) for men and 50 or 55 ($Agediff50_{it}$ and $Agediff55_{it}$) for women, are used as instruments for years in retirement. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors and having stayed for all three waves.

Table B.9: AB Estimation of Years in Retirement on Subjective and Physical Well-Being

	CES-D (0-30)		Depression (0,1)		Physical Problems	
	Men	Women	Men	Women	Men	Women
	(1)	(2)	(3)	(4)	(5)	(6)
retyrs	0.149	-0.090	0.008	0.002	-0.021	0.052
	(0.130)	(0.155)	(0.014)	(0.020)	(0.032)	(0.061)
L.rcesd10	-0.097	-0.084				
	(0.104)	(0.099)				
L.rdepress			0.057	0.077		
			(0.092)	(0.102)		
L.rphyhealth					-0.219	-0.150
					(0.294)	(0.232)
rage_sq	0.729	-0.330	0.056	-0.010	1.046***	0.285
	(0.545)	(0.724)	(0.052)	(0.069)	(0.297)	(0.368)
rmarried	-2.094*	-1.886*	-0.157*	0.016	-0.190	-0.146
	(1.210)	(1.076)	(0.094)	(0.033)	(0.854)	(0.389)
hchild	-0.682	0.089	0.007	0.006	0.051	-0.071
	(0.484)	(0.406)	(0.029)	(0.026)	(0.108)	(0.100)
hhhres	-0.018	-0.011	-0.012	-0.007	-0.059	-0.004
	(0.143)	(0.271)	(0.013)	(0.024)	(0.075)	(0.081)
lnhhadurbl	-0.134	0.035	-0.015	0.003	-0.047	0.012
	(0.100)	(0.127)	(0.010)	(0.010)	(0.047)	(0.044)
y2013	2.282	-0.840	0.172	-0.018	1.907***	0.162
	(1.390)	(1.754)	(0.127)	(0.170)	(0.653)	(0.764)
_cons	-23.037	22.063	-2.121	0.536	-40.663***	-7.344
	(23.906)	(28.686)	(2.221)	(2.705)	(12.364)	(13.988)
N	460	404	460	404	526	438
N_g	460	404	460	404	526	438
chi2	17.404	4.854	6.810	1.013	84.161	41.985

Note: Standard errors are clustered in individual level. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The sample is restricted to urban-residents living in the urban areas, having formally retired or working in the non-agricultural sectors and having stayed for all three waves.

Appendix C

Depression, Physical Health, Labour Market Exits and Entries of Older Informal Workers

C.1 Additional Figures

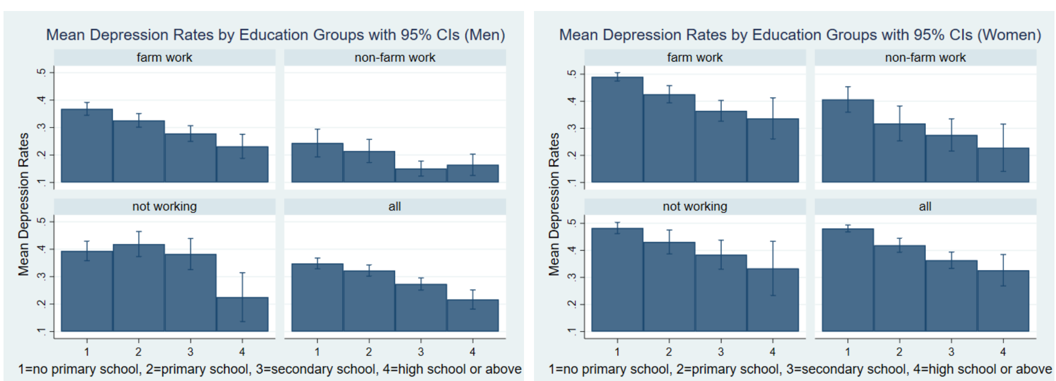

Figure C.1: Depression Rate over Educational Groups and labour Force Status

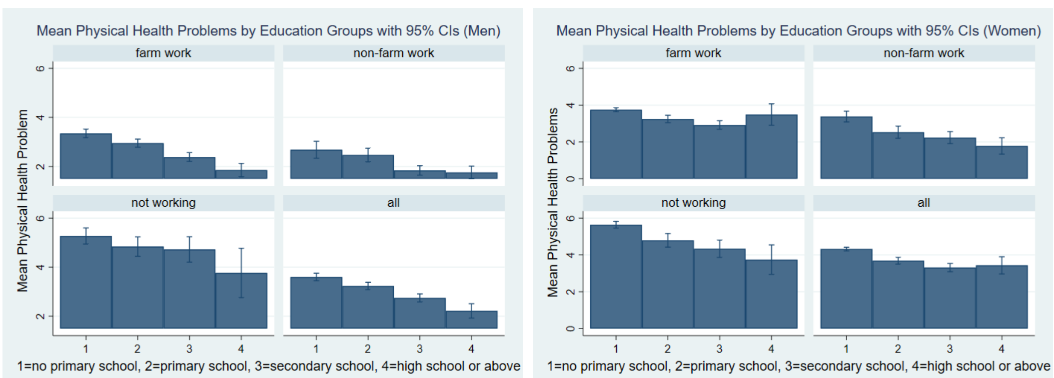

Figure C.2: Physical Health Problems over Educational Groups and labour Force Status

Bibliography

Adhvaryu, Achyuta, and Anant Nyshadham. 2017. "Health, enterprise, and labor complementarity in the household." *Journal of development economics*, 126: 91–111.

Alexopoulos, George S, Barnett S Meyers, Robert C Young, Scott Campbell, David Silbersweig, and Mary Charlson. 1997. "'Vascular depression'hypothesis." *Archives of general psychiatry*, 54(10): 915–922.

Andresen, Elena M, Judith A Malmgren, William B Carter, and Donald L Patrick. 1994. "Screening for depression in well older adults: Evaluation of a short form of the CES-D." *American journal of preventive medicine*, 10(2): 77–84.

Ardington, Cally, Anne Case, and Victoria Hosegood. 2009. "Labor supply responses to large social transfers: Longitudinal evidence from South Africa." *American economic journal: Applied economics*, 1(1): 22–48.

Arellano, Manuel, and Stephen Bond. 1991. "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations." *The review of economic studies*, 58(2): 277–297.

Arulampalam, Wiji, and Mark B Stewart. 2009. "Simplified implementation of the Heckman estimator of the dynamic probit model and a comparison with alternative estimators." *Oxford bulletin of economics and statistics*, 71(5): 659–681.

Atalay, Kadir, and Garry F Barrett. 2015. "The impact of age pension eligibility age on retirement and program dependence: Evidence from an Australian experiment." *Review of Economics and Statistics*, 97(1): 71–87.

Atalay, Kadir, Garry F Barrett, and Anita Staneva. 2019. "The effect of retirement on elderly cognitive functioning." *Journal of health economics*, 66: 37–53.

Attanasio, Orazio P, and Susann Rohwedder. 2003. "Pension wealth and household saving: Evidence from pension reforms in the United Kingdom." *American Economic Review*, 93(5): 1499–1521.

Baltagi, Badi. 2008. *Econometric analysis of panel data.* John Wiley & Sons.

Banks, James, and Fabrizio Mazzonna. 2012. "The effect of education on old age cognitive abilities: evidence from a regression discontinuity design." *The Economic Journal*, 122(560): 418–448.

Baranov, Victoria, Sonia Bhalotra, Pietro Biroli, and Joanna Maselko. 2020. "Maternal Depression, Women's Empowerment, and Parental Investment: Evidence from a Randomized Controlled Trial." *American Economic Review*, 110(3): 824–59.

Barry, Lisa C, Heather G Allore, Martha L Bruce, and Thomas M Gill. 2009. "Longitudinal association between depressive symptoms and disability burden among older persons." *Journals of Gerontology Series A: Biomedical Sciences and Medical Sciences*, 64(12): 1325–1332.

Battistin, Erich, Agar Brugiavini, Enrico Rettore, and Guglielmo Weber. 2009. "The retirement consumption puzzle: evidence from a regression discontinuity approach." *American Economic Review*, 99(5): 2209–26.

Beedie, Chris, Steven Mann, Robert Copeland, and Sarah Domone. 2014. "Identifying what works for local physical inactivity interventions."

Behncke, Stefanie. 2012. "Does retirement trigger ill health?" *Health economics*, 21(3): 282–300.

Bertrand, Marianne, Sendhil Mullainathan, and Douglas Miller. 2003. "Public policy and extended families: Evidence from pensions in South Africa." *the world bank economic review*, 17(1): 27–50.

Bingley, Paul, and Alessandro Martinello. 2013. "Mental retirement and schooling." *European Economic Review*, 63: 292–298.

Blundell, Richard, and Stephen Bond. 1998. "Initial conditions and moment restrictions in dynamic panel data models." *Journal of econometrics*, 87(1): 115–143.

Blundell, Richard W, Jack Britton, Monica Costa Dias, and Eric French. 2016. "The dynamic effects of health on the employment of older workers." *Michigan Retirement Research Center Research Paper*, , (2016-348).

Böckerman, Petri, and Pekka Ilmakunnas. 2009. "Unemployment and self-assessed health: evidence from panel data." *Health economics*, 18(2): 161–179.

Bonsang, Eric, Stéphane Adam, and Sergio Perelman. 2012. "Does retirement affect cognitive functioning?" *Journal of health economics*, 31(3): 490–501.

Börsch-Supan, Axel, and Morten Schuth. 2013. "Early retirement, mental health and social networks." *Active ageing and solidarity between generations in Europe: First results from SHARE after the economic crisis*, 337.

Bound, John, and Timothy Waidmann. 2007. "Estimating the health effects of retirement." *Michigan Retirement Research Center Research Paper No. UM WP*, 168.

Bound, John, Michael Schoenbaum, Todd R Stinebrickner, and Timothy Waidmann. 1999. "The dynamic effects of health on the labor force transitions of older workers." *Labour Economics*, 6(2): 179–202.

Bruce, Martha L. 2001. "Depression and disability in late life: directions for future research." *The American Journal of geriatric psychiatry*, 9(2): 102–112.

Butler, John S, and Robert Moffitt. 1982. "A computationally efficient quadrature procedure for the one-factor multinomial probit model." *Econometrica: Journal of the Econometric Society*, 761–764.

Case, Anne, and Christina Paxson. 2008. "Height, health, and cognitive function at older ages." *American economic review*, 98(2): 463–67.

Case, Anne, and Christina Paxson. 2009. "Early life health and cognitive function in old age." *American Economic Review*, 99(2): 104–09.

Celidoni, Martina, Chiara Dal Bianco, and Guglielmo Weber. 2017. "Retirement and cognitive decline. A longitudinal analysis using SHARE data." *Journal of health economics*, 56: 113–125.

Chamberlain, Gary. 1982. "Multivariate regression models for panel data." *Journal of econometrics*, 18(1): 5–46.

Chatterji, Pinka, Margarita Alegria, and David Takeuchi. 2011. "Psychiatric disorders and labor market outcomes: Evidence from the National Comorbidity Survey-Replication." *Journal of health economics*, 30(5): 858–868.

Cheng, Lingguo, Hong Liu, Ye Zhang, and Zhong Zhao. 2018*a*. "The health implications of social pensions: Evidence from China's new rural pension scheme." *Journal of Comparative Economics*, 46(1): 53–77.

Cheng, Lingguo, Hong Liu, Ye Zhang, and Zhong Zhao. 2018*b*. "The heterogeneous impact of pension income on elderly living arrangements: evidence from China's new rural pension scheme." *Journal of Population Economics*, 31(1): 155–192.

Chen, Huajuan, and Ada C Mui. 2014. "Factorial validity of the Center for Epidemiologic Studies Depression Scale short form in older population in China." *International psychogeriatrics*, 26(1): 49.

Chen, Simiao, Pascal Geldsetzer, and Till Bärnighausen. 2020. "The causal effect of retirement on stress in older adults in China: A regression discontinuity study." *SSM-population health*, 10: 100462.

Chen, Xi, Lipeng Hu, and Jody L Sindelar. 2020. "Leaving money on the table? Suboptimal enrollment in the new social pension program in China." *The Journal of the Economics of Ageing*, 15: 100233.

Chen, Xi, Tianyu Wang, and Susan H Busch. 2019. "Does money relieve depression? Evidence from social pension expansions in China." *Social Science & Medicine*, 220: 411–420.

Chen, Xi, Xiaobo Zhang, and Xin Zhang. 2017. "Smog in our brains: Gender differences in the impact of exposure to air pollution on cognitive performance."

Che, Yi, and Xin Li. 2018. "Retirement and health: Evidence from China." *China Economic Review*, 49: 84–95.

Clark, Andrew E, and Andrew J Oswald. 1994. "Unhappiness and unemployment." *The Economic Journal*, 104(424): 648–659.

Coe, Norma B, and Gema Zamarro. 2011. "Retirement effects on health in Europe." *Journal of health economics*, 30(1): 77–86.

Coe, Norma B, Hans-Martin von Gaudecker, Maarten Lindeboom, and Jürgen Maurer. 2012. "The effect of retirement on cognitive functioning." *Health economics*, 21(8): 913–927.

Coile, Courtney C. 2015. "ECONOMIC DETERMINANTS OF WORKERS'RETIREMENT DECISIONS." *Journal of Economic Surveys*, 29(4): 830–853.

Contoyannis, Paul, and Andrew M Jones. 2004. "Socio-economic status, health and lifestyle." *Journal of health economics*, 23(5): 965–995.

Contoyannis, Paul, and Jinhu Li. 2011. "The evolution of health outcomes from childhood to adolescence." *Journal of Health Economics*, 30(1): 11–32.

Contoyannis, Paul, Andrew M Jones, and Nigel Rice. 2004. "The dynamics of health in the British Household Panel Survey." *Journal of Applied Econometrics*, 19(4): 473–503.

Cunha, Flavio, and James Heckman. 2007. "The technology of skill formation." *American Economic Review*, 97(2): 31–47.

Currie, Janet. 2009. "Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development." *Journal of Economic Literature*, 47(1): 87–122.

Dalton, Patricio S, Sayantan Ghosal, and Anandi Mani. 2016. "Poverty and aspirations failure." *The Economic Journal*, 126(590): 165–188.

Dave, Dhaval, Inas Rashad, and Jasmina Spasojevic. 2006. "The effects of retirement on physical and mental health outcomes." National Bureau of Economic Research.

de Carvalho Filho, Irineu Evangelista. 2008. "Old-age benefits and retirement decisions of rural elderly in Brazil." *Journal of Development Economics*, 86(1): 129–146.

d'Hombres, Béatrice, Lorenzo Rocco, Marc Suhrcke, and Martin McKee. 2010. "Does social capital determine health? Evidence from eight transition countries." *Health economics*, 19(1): 56–74.

Disney, Richard, Carl Emmerson, and Matthew Wakefield. 2006. "Ill health and retirement in Britain: A panel data-based analysis." *Journal of health economics*, 25(4): 621–649.

Drew, Natalie, Edwige Faydi, Melvyn Freeman, and Michelle Funk. 2010. *Mental health and development: targeting people with mental health conditions as a vulnerable group.* World Health Organization.

Duflo, Esther. 2000. "Child health and household resources in South Africa: evidence from the old age pension program." *American Economic Review*, 90(2): 393–398.

Eibich, Peter. 2015. "Understanding the effect of retirement on health: Mechanisms and heterogeneity." *Journal of health economics*, 43: 1–12.

Fang, Hanming, and Jin Feng. 2018. "The Chinese Pension System." National Bureau of Economic Research.

Farré, Lídia, Francesco Fasani, and Hannes Mueller. 2018. "Feeling useless: the effect of unemployment on mental health in the Great Recession." *IZA Journal of Labor Economics*, 7(1): 8.

Feng, Jin, and Xiaohan Zhang. 2018. "Retirement and grandchild care in urban China." *Feminist Economics*, 24(2): 240–264.

Feng, Jin, Lixin He, and Hiroshi Sato. 2011. "Public pension and household saving: Evidence from urban China." *Journal of Comparative Economics*, 39(4): 470–485.

Feng, Jin, Qin Li, and James P Smith. 2020. "Retirement effect on health status and health behaviors in urban China." *World Development*, 126: 104702.

Ferrari, Alize J, Fiona J Charlson, Rosana E Norman, Scott B Patten, Greg Freedman, Christopher JL Murray, Theo Vos, and Harvey A Whiteford. 2013. "Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010." *PLoS med*, 10(11): e1001547.

French, Eric. 2005. "The effects of health, wealth, and wages on labour supply and retirement behaviour." *The Review of Economic Studies*, 72(2): 395–427.

French, Eric, and Jae Song. 2014. "The effect of disability insurance receipt on labor supply." *American economic Journal: economic policy*, 6(2): 291–337.

French, Eric, and John Bailey Jones. 2011. "The effects of health insurance and self-insurance on retirement behavior." *Econometrica*, 79(3): 693–732.

French, Eric, and John Jones. 2012. "Public pensions and labor supply over the life cycle." *International Tax and Public Finance*, 19(2): 268–287.

Frijters, Paul, David W Johnston, and Michael A Shields. 2014. "The effect of mental health on employment: evidence from Australian panel data." *Health economics*, 23(9): 1058–1071.

Fruehwirth, Jane Cooley, Sriya Iyer, and Anwen Zhang. 2019. "Religion and depression in adolescence." *Journal of Political Economy*, 127(3): 1178–1209.

Gale, William G. 1998. "The effects of pensions on household wealth: A reevaluation of theory and evidence." *Journal of Political Economy*, 106(4): 706–723.

García-Gómez, Pilar. 2011. "Institutions, health shocks and labour market outcomes across Europe." *Journal of health economics*, 30(1): 200–213.

García-Gómez, Pilar, Andrew M Jones, and Nigel Rice. 2010. "Health effects on labour market exits and entries." *Labour Economics*, 17(1): 62–76.

Garnaut, Ross, Ligang Song, and Yang Yao. 2006. "Impact and significance of state-owned enterprise restructuring in China." *The China Journal*, , (55): 35–63.

Garthwaite, Craig, Tal Gross, and Matthew J Notowidigdo. 2014. "Public health insurance, labor supply, and employment lock." *The Quarterly Journal of Economics*, 129(2): 653–696.

Giles, John, Albert Park, and Fang Cai. 2006. "How has economic restructuring affected China's urban workers?" *The China Quarterly*, 185: 61–95.

Glass, Thomas A, Carlos Mendes De Leon, Richard A Marottoli, and Lisa F Berkman. 1999. "Population based study of social and productive activities as predictors of survival among elderly Americans." *Bmj*, 319(7208): 478–483.

Graham, Carol. 2008. "Happiness and health: Lessons—and questions—for public policy." *Health affairs*, 27(1): 72–87.

Grip, Andries de, Maarten Lindeboom, and Raymond Montizaan. 2011. "Shattered dreams: the effects of changing the pension system late in the game." *The Economic Journal*, 122(559): 1–25.

Grossman, Michael. 1972. "On the concept of health capital and the demand for health." *Journal of Political economy*, 80(2): 223–255.

Gruber, Jonathan. 2000. "Disability insurance benefits and labor supply." *Journal of political economy*, 108(6): 1162–1183.

Gruber, Jonathan, and Brigitte C Madrian. 2002. "Health insurance, labor supply, and job mobility: A critical review of the literature." National Bureau of Economic Research.

Gruber, Jonathan, and David Wise. 1998. "Social security and retirement: An international comparison." *The American Economic Review*, 88(2): 158–163.

Gruber, Jonathan, and David Wise. 2005. "Social security programs and retirement around the world: Fiscal implications, introduction and summary." National Bureau of Economic Research.

Gustman, Alan L, and Thomas L Steinmeier. 1983. "A structural retirement model." National Bureau of Economic Research.

Gustman, Alan L, and Thomas L Steinmeier. 1994. "Retirement in a family context: A structural model for husbands and wives." National Bureau of Economic Research.

Gustman, Alan L, and Thomas L Steinmeier. 2005. "The social security early entitlement age in a structural model of retirement and wealth." *Journal of public Economics*, 89(2-3): 441–463.

Haan, Peter, and Michal Myck. 2009. "Dynamics of health and labor market risks." *Journal of health economics*, 28(6): 1116–1125.

Hanel, Barbara, and Regina T Riphahn. 2012. "The timing of retirement—New evidence from Swiss female workers." *Labour economics*, 19(5): 718–728.

Hauck, Katharina, and Nigel Rice. 2004. "A longitudinal analysis of mental health mobility in Britain." *Health economics*, 13(10): 981–1001.

Heckman, James J. 1981*a*. "Heterogeneity and state dependence." In *Studies in labor markets*. 91–140. University of Chicago Press.

Heckman, James J. 1981*b*. "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time% Discrete Data Stochastic Process, in Charles F. Manski and Daniel McFadden, eds., The Structural Analysis of Discrete Data." *MIT Press, Cambridge, MA*, 179: 195.

Hewlett, E, and V Moran. 2014. *OECD Health Policy Studies Making Mental Health Count The Social and Economic Costs of Neglecting Mental Health Care: The Social and Economic Costs of Neglecting Mental Health Care.* OECD Publishing.

Heyma, Arjan. 2004. "A structural dynamic analysis of retirement behaviour in the Netherlands." *Journal of Applied Econometrics*, 19(6): 739–759.

Huang, Chang-Quan, Bi-Rong Dong, Zhen-Chan Lu, Ji-Rong Yue, and Qing-Xiu Liu. 2010. "Chronic diseases and risk for depression in old age: a meta-analysis of published literature." *Ageing research reviews*, 9(2): 131–141.

Huang, Cheng, Michael R Phillips, Yali Zhang, Jingxuan Zhang, Qichang Shi, Zhiqiang Song, Zhijie Ding, Shutao Pang, and Reynaldo Martorell. 2013. "Malnutrition in early life and adult mental health: evidence from a natural experiment." *Social science & medicine*, 97: 259–266.

Huang, Wei, and Chuanchuan Zhang. 2021. "The power of social pensions: Evidence from China's new rural pension scheme." *American Economic Journal: Applied Economics*, 13(2): 179–205.

Huang, Wei, and Yi Zhou. 2013. "Effects of education on cognition at older ages: evidence from China's Great Famine." *Social Science & Medicine*, 98: 54–62.

Hultsch, David F, Christopher Hertzog, Brent J Small, and Roger A Dixon. 1999. "Use it or lose it: engaged lifestyle as a buffer of cognitive decline in aging?" *Psychology and aging*, 14(2): 245.

Hurd, Michael D. 1990. "The joint retirement decision of husbands and wives." In *Issues in the Economics of Aging*. 231–258. University of Chicago Press, 1990.

Hyslop, Dean R. 1999. "State dependence, serial correlation and heterogeneity in intertemporal labor force participation of married women." *Econometrica*, 67(6): 1255–1294.

Insler, Michael. 2014. "The health consequences of retirement." *Journal of Human Resources*, 49(1): 195–233.

Jensen, Robert T. 2004. "Do private transfers 'displace'the benefits of public transfers? Evidence from South Africa." *Journal of Public Economics*, 88(1-2): 89–112.

Kalwij, Adriaan, and Frederic Vermeulen. 2008. "Health and labour force participation of older people in Europe: what do objective health indicators add to the analysis?" *Health economics*, 17(5): 619–638.

Katon, Wayne, Elizabeth HB Lin, and Kurt Kroenke. 2007. "The association of depression and anxiety with medical symptom burden in patients with chronic medical illness." *General hospital psychiatry*, 29(2): 147–155.

Kaushal, Neeraj. 2014. "How public pension affects elderly labor supply and well-being: Evidence from India." *World Development*, 56: 214–225.

Kessler, Ronald C, and Evelyn J Bromet. 2013. "The epidemiology of depression across cultures." *Annual review of public health*, 34: 119–138.

Kofi Charles, Kerwin. 2004. "Is retirement depressing?: Labor force inactivity and psychological well-being in later life." In *Accounting for worker well-being*. 269–299. Emerald Group Publishing Limited.

Kolodziej, Ingo WK, and Pilar García-Gómez. 2019. "Saved by retirement: Beyond the mean effect on mental health." *Social Science & Medicine*, 225: 85–97.

Krueger, Alan B, and Bruce D Meyer. 2002. "Labor supply effects of social insurance." *Handbook of public economics*, 4: 2327–2392.

Lei, Xiaoyan, and Hong Liu. 2018. "Gender difference in the impact of retirement on cognitive abilities: Evidence from urban China." *Journal of Comparative Economics*, 46(4): 1425–1446.

Lei, Xiaoyan, and Wanchuan Lin. 2009. "The new cooperative medical scheme in rural China: Does more coverage mean more service and better health?" *Health economics*, 18(S2): S25–S46.

Lei, Xiaoyan, Xiaoting Sun, John Strauss, Peng Zhang, and Yaohui Zhao. 2014*a*. "Depressive symptoms and SES among the mid-aged and elderly in China: Evidence from the China Health and Retirement Longitudinal Study national baseline." *Social Science & Medicine*, 120: 224–232.

Lei, Xiaoyan, Xiaoting Sun, John Strauss, Yaohui Zhao, Gonghuan Yang, Perry Hu, Yisong Hu, and Xiangjun Yin. 2014*b*. "Health outcomes and socio-economic status among the mid-aged and elderly in China: Evidence from the CHARLS national baseline data." *The Journal of the Economics of Ageing*, 3: 29–43.

Lei, Xiaoyan, Yan Shen, James P Smith, and Guangsu Zhou. 2015. "Do social networks improve Chinese adults' subjective well-being?" *The Journal of the Economics of Ageing*, 6: 57–67.

Li, Qin, Yafeng Wang, and Yaohui Zhao. 2018. "The Impact of China's New Rural Pension Program on Elderly Labor, Grandchild Care, and Old-Age Support." *Feminist Economics*, 24(2): 265–287.

Li, Wentian. 2017. "An empirical study on the health of retirement to the working population—based on China Health and Retirement Longitudinal Study." *American Journal of Industrial and Business Management*, 7(4): 429–443.

Llena-Nozal, Ana, Maarten Lindeboom, and France Portrait. 2004. "The effect of work on mental health: does occupation matter?" *Health economics*, 13(10): 1045–1062.

Luo, Meng Sha, Ernest Wing Tak Chui, and Lydia W Li. 2019. "The longitudinal associations between physical health and mental health among older adults." *Aging & mental health*, 1–9.

Mastrobuoni, Giovanni. 2009. "Labor supply effects of the recent social security benefit cuts: Empirical estimates using cohort discontinuities." *Journal of public Economics*, 93(11-12): 1224–1233.

Mazzonna, Fabrizio, and Franco Peracchi. 2012. "Ageing, cognitive abilities and retirement." *European Economic Review*, 56(4): 691–710.

Mazzonna, Fabrizio, and Franco Peracchi. 2017. "Unhealthy retirement?" *Journal of Human Resources*, 52(1): 128–151.

Meng, Xin, and Sen Xue. 2020. "Social networks and mental health outcomes: Chinese rural–urban migrant experience." *Journal of Population Economics*, 33(1): 155–195.

Messe, Pierre-Jean, and François-Charles Wolff. 2019. "Healthier when retiring earlier? Evidence from France." *Applied Economics*, 1–22.

Moral-Benito, Enrique. 2013. "Likelihood-based estimation of dynamic panels with predetermined regressors." *Journal of Business & Economic Statistics*, 31(4): 451–472.

Moral-Benito, Enrique, Paul Allison, and Richard Williams. 2019. "Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond." *Applied Economics*, 51(20): 2221–2232.

Mundlak, Yair. 1978. "On the pooling of time series and cross section data." *Econometrica: journal of the Econometric Society*, 69–85.

Neuman, Kevin. 2008. "Quit your job and get healthier? The effect of retirement on health." *Journal of Labor Research*, 29(2): 177–201.

Nickell, Stephen. 1981. "Biases in dynamic models with fixed effects." *Econometrica: Journal of the Econometric Society*, 1417–1426.

of Statistics of China, National Bureau. 2018. *China statistical yearbook 2018.* Javvin Technologies Incorporated.

Ohrnberger, Julius, Eleonora Fichera, and Matt Sutton. 2017*a*. "The dynamics of physical and mental health in the older population." *The Journal of the Economics of Ageing*, 9: 52–62.

Ohrnberger, Julius, Eleonora Fichera, and Matt Sutton. 2017*b*. "The relationship between physical and mental health: A mediation analysis." *Social Science & Medicine*, 195: 42–49.

Posel, Dorrit, James A Fairburn, and Frances Lund. 2006. "Labour migration and households: A reconsideration of the effects of the social pension on labour supply in South Africa." *Economic modelling*, 23(5): 836–853.

Radloff, Lenore Sawyer. 1977. "The CES-D scale: A self-report depression scale for research in the general population." *Applied psychological measurement*, 1(3): 385–401.

Ranchhod, Vimal. 2006. "The effect of the South African old age pension on labour supply of the elderly." *South African Journal of Economics*, 74(4): 725–744.

Ravesteijn, Bastian, Hans van Kippersluis, and Eddy van Doorslaer. 2013. "The contribution of occupation to health inequality." In *Health and inequality*. 311–332. Emerald Group Publishing Limited.

Richards, Marcus, Beverly Shipley, Rebecca Fuhrer, and Michael EJ Wadsworth. 2004. "Cognitive ability in childhood and cognitive decline in mid-life: longitudinal birth cohort study." *Bmj*, 328(7439): 552.

Rohwedder, Susann, and Robert J Willis. 2010. "Mental retirement." *Journal of Economic Perspectives*, 24(1): 119–38.

Roodman, David. 2011. "Fitting fully observed recursive mixed-process models with cmp." *The Stata Journal*, 11(2): 159–206.

Ruhm, Christopher J. 1991. "Are workers permanently scarred by job displacements?" *The American economic review*, 81(1): 319–324.

Rust, John, and Christopher Phelan. 1997. "How social security and medicare affect retirement behavior in a world of incomplete markets." *Econometrica: Journal of the Econometric Society*, 781–831.

Salthouse, Timothy A. 1996. "The processing-speed theory of adult age differences in cognition." *Psychological review*, 103(3): 403.

Semykina, Anastasia, and Jeffrey M Wooldridge. 2018. "Binary response panel data models with sample selection and self-selection." *Journal of Applied Econometrics*, 33(2): 179–197.

Shai, Ori. 2018. "Is retirement good for men's health? Evidence using a change in the retirement age in Israel." *Journal of health economics*, 57: 15–30.

Shu, Lei. 2018. "The effect of the New Rural Social Pension Insurance program on the retirement and labor supply decision in China." *The Journal of the Economics of Ageing*, 12: 135–150.

Skoog, Ingmar, and Deborah Gustafson. 2006. "Update on hypertension and Alzheimer's disease." *Neurological research*, 28(6): 605–611.

Smith, James P, John J McArdle, and Robert Willis. 2010. "Financial decision making and cognition in a family context." *The Economic Journal*, 120(548): F363–F380.

Sobocki, Patrik, Bengt Jönsson, Jules Angst, and Clas Rehnberg. 2006. "Cost of depression in Europe." *Journal of Mental Health Policy and Economics.*

Stern, Yaakov. 2002. "What is cognitive reserve? Theory and research application of the reserve concept."

Stewart, Jennifer M. 2001. "The impact of health status on the duration of unemployment spells and the implications for studies of the impact of unemployment on health status." *Journal of health economics*, 20(5): 781–796.

Stewart, Mark B. 2007. "The interrelated dynamics of unemployment and low-wage employment." *Journal of applied econometrics*, 22(3): 511–531.

Stewart, Walter F, Judith A Ricci, Elsbeth Chee, Steven R Hahn, and David Morganstein. 2003. "Cost of lost productive work time among US workers with depression." *Jama*, 289(23): 3135–3144.

Stine-Morrow, Elizabeth AL. 2007. "The Dumbledore hypothesis of cognitive aging." *Current Directions in Psychological Science*, 16(6): 295–299.

Sun, Jessica Ya. 2020. "Welfare consequences of access to health insurance for rural households: Evidence from the New Cooperative Medical Scheme in China." *Health Economics*, 29(3): 337–352.

Taylor, Warren D, Howard J Aizenstein, and George S Alexopoulos. 2013. "The vascular depression hypothesis: mechanisms linking vascular disease with depression." *Molecular psychiatry*, 18(9): 963–974.

Trevisan, Elisabetta, and Francesca Zantomio. 2016. "The impact of acute health shocks on the labour supply of older workers: Evidence from sixteen European countries." *Labour Economics*, 43: 171–185.

United Nations. 2019. "World population prospects 2019: department of economic and social Affairs." *World Population Prospects 2019.*

Van der Klaauw, Wilbert, and Kenneth I Wolpin. 2008. "Social security and the retirement and savings behavior of low-income households." *Journal of econometrics*, 145(1-2): 21–42.

Verbeek, Marno, and Theo Nijman. 1992. "Testing for selectivity bias in panel data models." *International Economic Review*, 681–703.

Vos, Theo, Abraham D Flaxman, Mohsen Naghavi, Rafael Lozano, Catherine Michaud, Majid Ezzati, Kenji Shibuya, Joshua A Salomon, Safa Abdalla, Victor Aboyans, et al. 2012. "Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010." *The lancet*, 380(9859): 2163–2196.

Wang, Mo. 2007. "Profiling retirees in the retirement transition and adjustment process: Examining the longitudinal change patterns of retirees' psychological well-being." *Journal of Applied Psychology*, 92(2): 455.

Wang, Mo, and Junqi Shi. 2014. "Psychological research on retirement." *Annual review of psychology*, 65: 209–233.

Wilson, Robert S, LL Barnes, CF Mendes De Leon, NT Aggarwal, JS Schneider, J Bach, J Pilat, Laurel A Beckett, SE Arnold, DA Evans, et al. 2002. "Depressive symptoms, cognitive decline, and risk of AD in older persons." *Neurology*, 59(3): 364–370.

Windmeijer, Frank. 2005. "A finite sample correction for the variance of linear efficient two-step GMM estimators." *Journal of econometrics*, 126(1): 25–51.

Wooldridge, Jeffrey M. 2005. "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity." *Journal of applied econometrics*, 20(1): 39–54.

Wooldridge, Jeffrey M. 2010. *Econometric analysis of cross section and panel data.* MIT press.

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

Zantinge, Else M, Matthijs van den Berg, Henriette A Smit, and H Susan J Picavet. 2013. "Retirement and a healthy lifestyle: opportunity or pitfall? A narrative review of the literature." *The European Journal of Public Health*, 24(3): 433–439.

Zhang, Wenjuan, and Dan Tang. 2008. "The new rural social pension insurance programme of Baoji City." *HelpAge International-Asia/Pacific, Chiang Mai.*

Zhang, Yi, Martin Salm, and Arthur van Soest. 2018. "The effect of retirement on healthcare utilization: Evidence from China." *Journal of health economics*, 62: 165–177.

Zhao, Qing, Zhen Li, and Taichang Chen. 2016. “The impact of public pension on household consumption: Evidence from China’s survey data.” *Sustainability*, 8(9): 890.

Zhao, Rui, and Yaohui Zhao. 2018. “The gender pension gap in China.” *Feminist Economics*, 24(2): 218–239.