Available Online: 4 November 2026 DOI: 10.54254/2977-5701/2026.38672

The impact of retirement on cognitive function of urban women under China's age-50 retirement policy: a fuzzy regression discontinuity study

Lai Wei

University of Michigan, Ann Arbor, USA

Weilai931@163.com

Abstract. Existing evidence on retirement and cognitive function is inconclusive. Thus, using five national waves (2011-2020) of the China Health and Retirement Longitudinal Study (CHARLS), this study examines urban female employees aged 45-65 covered by China's age-50 statutory retirement policy. A fuzzy regression discontinuity design uses the age-50 as an instrument for retirement. Accordingly, the identified Local Average Treatment Effect (LATE) applies primarily to urban women subject to the age-50 retirement rule, mainly blue-collar or non-managerial workers, rather than all urban women in China. At the cutoff, no statistically significant immediate net effect of retirement is observed on either mental intactness or episodic memory. Donut regression excluding individuals aged 50 shows a significant negative effect of retirement on mental intactness. This is consistent with a potential delayed manifestation. Therefore, a zero immediate net effect does not mean that retirement is entirely unrelated to cognition. However, because this study employs a pooled cross-sectional regression discontinuity design and does not track changes in the same individuals over time, these results should be interpreted cautiously.

Keywords: retirement, cognitive function, Chinese women, age-50 retirement policy, fuzzy regression discontinuity

1. Introduction

According to the United Nations Development Programme[1], China has the world's largest and fastest-growing elderly population. In addition, according to China's National Bureau of Statistics, until 2025, China's old-age dependency ratio had reached 23.1%. Against this background, cognitive health has become an important factor in aging area. This is because cognitive function is a determinant of older people's life quality,

[1] United Nations Development Programme. (2025, December 22). A measurement study on multidimensional vulnerabilities of the elderly population in China. https://www.undp.org/china/publications/measurement-study-multidimensional-vulnerabilities-elderly-population-china. This report is a UNDP publication and is provided here solely as a source of background data on China's elderly population.

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

independent living ability. Moreover, there are gendered differences in the process of cognitive aging. Although women have longer average life expectancy, more and more evidences suggest that the prevalence of Alzheimer's disease is higher among women than among men, making cognitive health attract greater attention.

Retirement is the major turning point in women's life. It not only indicates a break in labor participation, which changes economic income, but also represents a shift in daily routines and social interaction. In theory, this transition may affect cognitive function in multiple ways, including a reduction in cognitive stimulation, a narrowing of social circle, and changes in daily routines. However, there is still no consensus as to whether these mechanisms could translate into observable causal effects. Studies based on different countries, different data sources, and different causal inference strategies have reported positive, negative, and null effects of retirement on cognition [1]. This ambiguity also stems from methodological challenges. Retirement decisions are influenced by a wide range of factors, including macroeconomic policies, personal health, wealth accumulation, and life satisfaction. Most of these may have an impact on one's cognitive function. Thus, traditional regression methods cannot separate the effects of retirement from those of other factors.

The retirement policy in China makes this problem more complex. China has two different statutory retirement age for women. One statutory retirement age for women in Blue-collar is 50, whereas for women in white-collar it is 55. In practice, however, retirement time may also vary according to individual preference. Blue-collar workers primarily involve physical labor. Compared to white-collar workers, they worked in harsher environments and require more physical exertion. Long-term engagement in this type of work may lead to accelerated physical wear and tear. The statutory retirement age of 50 provides a systemic exit opportunity for those engaged in long-term, high-intensity physical labor. As a result, the vast majority of individuals retire at that age, leading to a high level of compliance with the system. In contrast, white-collar jobs primarily involve mental labor, the decline in health capital occurs at a slower rate. As a result, the mandatory retirement age of 55 is applied relatively flexibly among white-collar workers, and, combined with the widespread practice of rehiring, this flexibility has led to a more dispersed pattern of retirement around this age.

Besides, there are differences in pension coverage and retirement systems between urban and rural areas, as well as between public and private sector employees. Moreover, most of Chinese women become the major caregivers in their family after retirement. Responsibilities such as grandchild care and household chores fall disproportionately on women. This shift may provide alternative cognitive stimulation after retirement, yet it can also exacerbate social isolation. Therefore, research based on Chinese data that uses rigorous causal identification methods and focuses specifically on cognitive function among women is of considerable academic value.

In line with this research direction, this study uses causal identification framework to estimate the effect of retirement on cognitive function among urban women subject to the age-50 retirement policy in China.

This study employs a fuzzy regression discontinuity design, using the statutory retirement age for women in China as an exogenous institutional boundary to create identification conditions that approximate a randomized experiment near the cutoff. The data are from the China Health and Retirement Longitudinal Study (CHARLS), specifically the 2011 baseline survey and four follow-up surveys (2013, 2015, 2018, and 2020). These five survey waves form a pooled cross-sectional dataset. The population used in this study was limited to women aged from 45 to 65 who are enrolled in either the pension insurance program for government agencies and public institutions or the pension insurance program for enterprise employees. Cognitive function was divided into mental intactness and episodic memory, which were analyzed separately.

2. Literature review

Population aging has become a central trend in global demographic shifts, and health issues among the elderly have consequently emerged as a key focus of aging research. Against a background of increasingly refined health research, cognitive health has gradually come to be recognized as a key area within aging research owing to its critical impact on older adults' quality of life and ability to live independently. A large amount of research shows that physical function [2], anthropometric measures [3], and daily activity capacity [4] all have an impact on cognitive performance of women. Gendered differences in women's cognitive health and trajectory of cognitive decline are receiving increasing attention. Biological factors such as sex-specific genetic differences [5], changes in hormone levels [6], and reproductive characteristics [7] may all play a role in the patterns of women's cognitive decline. With regard to gendered differences in cognitive aging, women exhibit more pronounced declines in verbal episodic memory, associative memory, and short-term memory, while men show more marked deterioration in spatial relationship memory, comprehension and recognition, and visual structural abilities [8].

The gendered differences stem mainly from biological and physiological mechanisms. However, changes in acquired social roles may also play a key role in the women's cognitive aging progress. Retirement is one of the key factors as a turning event marking the end of labor participation.

The basis for the impact of retirement on cognitive function can be traced back to the psychological retirement hypothesis. Rohwedder and Willis [9] proposed two core mechanisms about this hypothesis. First, retirement shifts individuals' daily environment from high cognitive stimulation to low cognitive stimulation. Before retirement, working space provided structured cognitive challenges, including information processing, multitasking coordination, and more. However, after retirement, these stimulation decrease, and daily life tends to become monotonous. Second, even before officially retirement, people often reduce their psychological investment in their work when they anticipate retirement. This generates the effect known as "on-the-job retirement" effect, which causes cognitive function to begin declining even before retirement. Moreover, workplace is not only a space for economic activity but also a primary venue for social interaction. A study by Nie et al. [10] showed that older adults with broader social networks and greater participation in social activities have better cognitive function. Retirement cuts individuals off from their former workplace social circles, leading to a decrease in the frequency of social interactions. As a result, the social stimulation that used to help maintain cognitive activity is reduced, which may further accelerate the decline in cognitive function.

It is worth noting that cognitive trajectories after retirement are not homogeneous but may vary significantly depending on an individual's occupation and gender. Women, individuals with higher socioeconomic status, those in good health prior to retirement, and those who were physically active before retirement experience more significant cognitive benefits from retirement [11]. In terms of job type, older adults engaged in more complex work have better cognitive function and a lower risk of developing dementia [12]. Greenberg and Burgard [13] further pointed out that cognitive function in elder persons is associated with the intensity and complexity of work experienced throughout their lives, and also varies by gender. Regarding to health status, Xue et al. [14] found that, compared with women who retired for no specific reason at the Standard Retirement Age (SPA), women who retired due to poor health condition experienced a more rapid decline in memory over time, and this association was primarily observed among older women in the highest occupational status.

To empirically verify the causal effects involved, the endogeneity challenges inherent in life-course research must be addressed. An individual retirement is often intertwined with factors such as health condition and economic circumstances. These factors also affect cognitive function, leading to significant self-selection

bias in traditional regression estimates. Schwartz and Glymour [15] proposed that this can be offset through nested structural models, longitudinal instrumental variable methods, and diverse causal inference techniques. This methodological approach has received considerable attention in the cognitive function research. Consequently, a large number of studies began to adopt quasi-experimental designs in an effort to examine the causal relationship between retirement and cognition. However, research has not yet reached a consensus. The relationship between retirement and women's cognitive function is not fixed but varies significantly depending on the causal identification strategies employed, the data sources used, and sample selection methods.

Specifically, Hale et al. [16] employed a parameterized g-formula to estimate the counterfactual effects of delaying retirement until age 67. They found that delaying retirement effectively prevents the cognitive decline associated with retirement. Shin and Fan [17], on the other hand, using external health shocks serving as the instrumental variables, found that the decline in fluid intelligence scores resulting from lower employment rates was observed only among older women. Sato et al. [11] applied the instrumental variable causal forest method and showed that retirees scored higher on word recall tests than working-age adults and exhibited superior cognitive functioning. Chen [18] employed a fuzzy regression discontinuity design and found that a spouse's retirement has a positive impact on both the physical and mental health of women. In addition, a study using data from the China Family Panel Studies (CFPS) employed a fuzzy regression discontinuity design to analyze the effects of a spouse's retirement on an individual's cognitive health [19]. However, the analysis found that a spouse's retirement has a significant negative impact on an individual's cognitive health.

In addition to studies finding that retirement may have cognitive benefits or drawbacks, some studies have found that retirement has no significant effect on cognitive function. A study by Wang et al. [20] employed a fuzzy regression discontinuity design to examine the relationship between retirement and physical and mental health. They found that retirement does not have a significant impact on health indicators, including depression and cognitive function.

In addition to examining the cognitive consequences of retirement itself, other studies have used causal inference methods to further explore measures that may lead to cognitive benefits after retirement. Lee et al. [21] used a quasi-experimental design combined with the difference-in-differences method to evaluate the impact of a digital literacy education program on the cognitive function of rural older adults. They found that the program significantly improved participants' smartphone proficiency and cognitive function. Kim et al. [22] used a difference-in-differences method, and found that re-entering the labor market in later life has a protective effect on cognitive function. The combination of matching and the difference-in-differences method is common and crucial. Matching mitigates selection bias in observable covariates, while the difference-in-differences method reduces the impact of unobservable differences that do not vary across time. This dual-defense strategy significantly enhances the reliability of causal inferences.

Based on the above studies, current causal identification strategies in this field can be categorized into three main approaches. First, longitudinal comparison strategies, such as using the difference-in-differences method in combination with propensity score matching; second, strategies for addressing endogeneity, such as instrumental variables method; and third, quasi-randomized experimental strategies, such as the regression discontinuity design. With regard to the first category of methods, the difference-in-differences method is often used with propensity score matching [21, 22]. The former reduces sample selection bias by controlling for observable covariates, while the latter further eliminates unobservable differences between the treatment and control groups which do not vary over time. Thereby they can create a dual defense mechanism to enhance the credibility of causal identification. In the second category of methods, the instrumental variables approach uses exogenous variables to identify the causal effects of employment and other behaviors [17]; its core lies in addressing unobservable confounding and reverse causality issues. The third type of method, the

regression discontinuity design, uses the exogenous boundary of the policy to construct a quasi-randomized experiment [18, 19]. By treating retirement status as approximately randomly assigned around the cutoff, this approach compares individuals on either side of the threshold, who are difficult to manipulate precisely, to effectively control for confounding factors that do not jump at the cutoff along with the assignment variable. Thus, in the absence of randomized experiments, it provides credible causal inferences for identifying the cognitive effects of policy shocks.

Based on the above literature review, the innovative aspects of this study are primarily reflected in the following three areas. First, in terms of data, this study uses the data from all five waves of the China Health and Retirement Longitudinal Study (CHARLS) conducted between 2011 and 2020. These data cover longer time span, which provides a more comprehensive snapshot of the cross-sectional distribution of cognitive function among individuals around retirement age. Second, regarding to the research methodology, this study employs a fuzzy regression discontinuity design to address the endogeneity issue in retirement. Whether an individual retires is influenced by a variety of factors, which also affect cognitive function, leading to significant self-selection bias in traditional estimates. This study uses the statutory retirement age as an exogenous institutional instrumental variable to enhance the reliability of causal identification. Third, in terms of research strategy, this study adopts a more cautious approach in sample, cutoff and model selection. This paper strictly limits the sample to women who have reached the statutory retirement age, excluding those to whom the pension system does not apply or for whom retirement policies are unclear. At the same time, by selecting cutoffs based on the F-statistic from the first stage and graphical evidence, it avoids arbitrary cutoff selection and enhance the rigor of the procedure.

3. Institutional policy background

According to China's retirement policies, as of 2020, the statutory retirement age for women in blue-collar jobs is 50, while for women in white-collar jobs, it is 55. However, due to disparities in employment and household registration status, the statutory retirement age and a stable pension are typically reserved for the urban resident population with formal employment. Data from the 2008 pilot survey of the China Health and Retirement Longitudinal Study (CHARLS) show that employment rates among rural residents in China remain relatively high even after they reach the mandatory retirement age typically applied to urban residents [23].

In addition, China's pension system encompasses a variety of types. Until 2009, rural residents, the urban unemployed, homemakers, who are predominantly women, and rural migrant workers were still excluded from the social security system. The implementation of the Urban and Rural Residents' Social Pension (URRSP) covered the aforementioned groups as well as urban residents excluded from other pension systems [24]. These groups generally have no mandatory retirement or exit from the labor market provisions.

In summary, this study limits its sample to urban female respondents who participate in the pension insurance programs for government agencies and public institutions, as well as enterprise employees.

4. Data sources and study population

The data analyzed in this study are drawn from the China Health and Retirement Longitudinal Study (CHARLS). CHARLS is the first nationally representative survey of the middle-aged and older population. The national baseline survey was conducted in 2011, followed by four rounds of regular follow-up surveys in 2013, 2015, 2018, and 2020. The CHARLS sample represents the resident population aged 45 and older and was selected using a stratified, multistage PPS random sampling strategy. The CHARLS questionnaire primarily consists of the following modules: basic information, family information, health status and

functioning, work and retirement, household income and expenditures, personal income, and the COVID-19 pandemic [25][2].

This study uses data from the 2011 CHARLS national baseline survey and the four rounds of national follow-up surveys conducted from 2013 to 2020, focusing on urban female respondents aged 45 to 65 who are enrolled in the pension insurance programs for government agencies and public institutions or for enterprise employees. This study initially selected 3,804 observations.

The following data filtering steps were as follows: first, excluding individuals who had never worked (6 observations); second, excluding individuals with completely missing cognitive function variables or missing retirement status (287 observations). The final valid sample included in the analysis consisted of 3,416 observations.

Whether the participants in this study are retired is determined by whether they have completed the retirement procedures, as defined by the CHARLS survey variable "Retirement Processed".

5. Cognitive function

This study analyzes the impact of retirement on different aspects of cognitive function of the elderly with different metrics of cognitive function. Based on previous studies [26-28], cognitive function is divided into two modules, mental intactness and episodic memory. The module of mental intactness is assessed by the Telephone Interview Cognitive Status (TICS) and visuospatial test of intelligence and reasoning. The raw score ranges from 0 to 11.

Episodic memory is composed of immediate and delayed memory. The raw score for the episodic memory module is derived by averaging the immediate and delayed memory, and is scored between 0 and 10.

As mentioned by Meng et al. [29], a combination of these tests above have been recognized to be validated for elderly people in China.

However, the raw scores for the two modules are typical count variables with distinctly right-skewed distributions. When used directly as dependent variables in local linear regressions, the coefficients of the local linear regressions cannot be compared across different studies. Therefore, this paper first transformed the raw scores into Z-scores [30]. Thus, the dependent variable is standardized and expressed as a continuous variable, with a mean of 0 and standard deviation of 1.

It needs to be shown that this transformation will not affect the causal identification of regression discontinuity designs. Lee and Lemieux [31] noted that, applying a z-transformation is an innocuous linear transformation, and would have no effect on regression coefficients in a regression with a cutoff. Identifying treatment effects requires a difference in conditional expectations at the cutoff. Because of this, standardized regression coefficients have a clear causal interpretation and enhance interpretation.

6. Control variables

This paper uses several sociodemographic variables as control variables. These include residence status, educational level, and marital status. In this paper, educational attainment is categorized as "low" and "high"

[2] Zhao, Y., Strauss, J., Yang, G., Giles, J., Hu, P., Hu, Y., Lei, X., Liu, M., Park, A., Smith, J. P., & Wang, Y. (2013). China Health and Retirement Longitudinal Study: 2011–2012 national baseline user's guide. National School of Development, Peking University. This user manual provides instructions on the CHARLS questionnaire modules and technical details of the baseline survey.

based on whether the respondents' highest level of education falls within the scope of China's nine-year compulsory education. Residence status is categorized as agricultural, non-agricultural, unified, or no HuKou. Marital status is categorized as having a spouse and not having a spouse. In addition, the data used in this study cover five waves of the CHARLS survey, and the same respondent may appear in different survey waves. Therefore, this study includes the survey waves as control variables in the model to account for time-fixed effects.

7. Research design

This paper combines data from the five waves of the CHARLS survey into a pooled cross-sectional dataset and employs a fuzzy regression discontinuity design to examine the effects of retirement on cognitive function among urban women subject to the age-50 retirement policy. The research model for this paper is as follows (Equation (1)):

$$Cog_{it} = \beta_0 + \beta_1 \mathrm{Ret}_{it} + \beta_2 X_{it} + f(Age_{it}) + \tau_t + \varepsilon_{it} \quad (1)$$

In this model, Cog_{it} is the outcome variable, which measures the cognitive function score of female sample individual i at time t. Retit is a binary treatment variable that indicates the individual's retirement status. In this paper, this binary variable is defined using the "Retirement Processed" variable from the CHARLS survey. $f(Ageit)$ represents the age smoothing function, which is estimated using a local polynomial fit and optimized bandwidth selection. X_{it} refers to individual characteristic variables, including residence status, educational level, and marital status, and the survey wave to which the sample belongs.

Since this study uses combined data from the CHARLS National Survey, the same respondent may appear multiple times across different years. This may result in serial correlation within individuals. This study employs cluster-robust standard errors at the individual level in all regression estimates.

Furthermore, retirement typically occurs only once within the age 45-65, so this study did not employ a strategy combining individual fixed effects with regression discontinuity. The results of estimates incorporating individual fixed effects will be reported as a robustness test.

8. Validation test

The following two conditions must be met for the validation of the fuzzy regression discontinuity design: First, there must be a discontinuity in the conditional probability at the cutoff. Second, in the absence of treatment, the outcome variable for individuals should not undergo systematic jumps on either side of the cutoff but should instead vary smoothly with the treatment variable [32].

To test Condition 1, this study uses the actual age minus the statutory retirement age as the standardized independent variable and conducts first-stage regression analyses using 50 and 55 as cutoffs.

Figures 1 and 2 illustrate the changes in the probability of retirement at different cutoffs. The results show that at the age of 50, the probability of retirement has a obvious jump of approximately 0.7 near the cutoff. At the age of 55, the jump in the probability of retirement is slightly smaller, at about 0.4. This suggests that, compared to age 55, age 50 is a more significant institutional cutoff and women's retirement behavior tends to be more concentrated at this age.

Table 1 reports the results of first-stage regression. When evaluating the validity of the instrumental variables in the fuzzy regression discontinuity design, the first-stage F-statistic is a key diagnostic indicator. It is generally accepted that if the F-statistic for this stage is less than 10, it indicates a weak correlation between

the instrumental variable and the endogenous variable. This would lead to significant bias in subsequent causal estimates [33, 34].

When using age 50 as the cutoff, the F-statistic is 404.11, far beyond the critical value of 10. This strongly rejects the weak instrumental variable hypothesis and confirms the existence of a statistically jump in retirement status at age 50.

In contrast, when using age 55 as the cutoff, the F-statistic is 7.43, falling below the critical value of 10. Therefore, this is an example of a weak instrumental variable. This reflects the flexibility in retirement policy implementation for white-collar workers and the presence of re-employment phenomena.

A comparison of the data in Figure 1, Figure 2 and Table 1 reveals that the probability of retirement for women appears a more pronounced discontinuity at age 50. Therefore, to ensure the reliability of the model's results, this study selects age 50 as the analytical cutoff for the fuzzy regression discontinuity design.

It should be noted that, by using age 50 as the sole cutoff, this study identifies the LATE for women whose retirement behavior is affected by the age-50 statutory retirement rule. This group consists mainly of blue-collar and non-managerial urban female employees. The estimates do not identify the effect for women subject to the age-55 rule, nor for all urban women in China. The reason this paper does not categorize participants by educational background or occupation in its empirical analysis is that the CHARLS dataset contains a significant number of missing values for variables used to identify an individual's occupational status (e.g., cadre vs. worker), and there is no deterministic correspondence between educational background, place of employment, and actual retirement age. Highly educated women may also work in jobs subject to a retirement age of 50, and there are support staff positions in government agencies and public institutions also require retirement at age 50. Therefore, this paper defines the scope of the LATE and uses full-sample covariate control to account for potential heterogeneity.

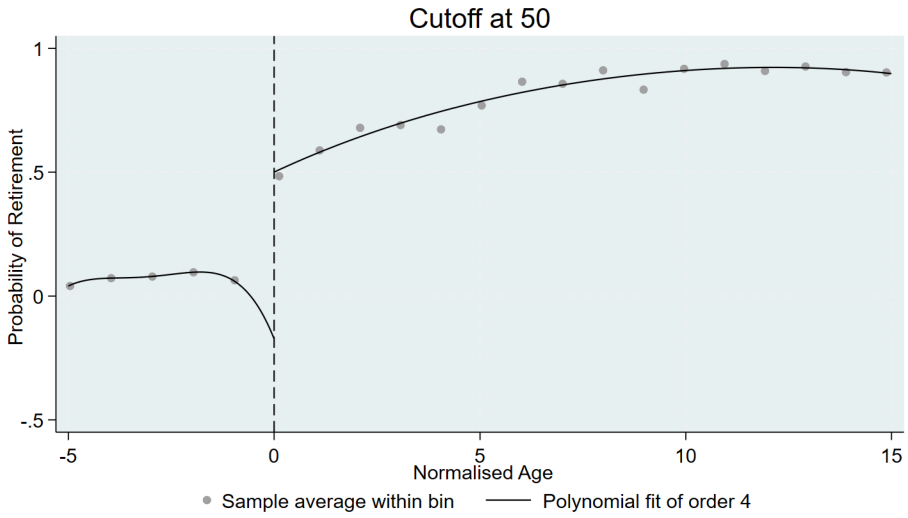


Figure 1. Retirement situation for women at the age of 50 as the cutoff point

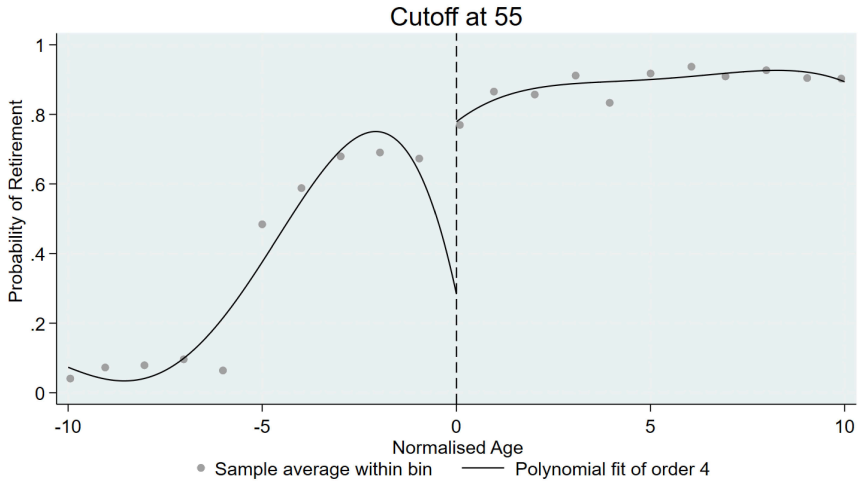


Figure 2. Retirement situation for women at the age of 55 as the cutoff point

Table 1. First-stage regression at age 50 and 55

	Cutoff = 50	Cutoff = 55
Treated	0.460***	0.078***
	(0.023)	(0.029)
Centered Age	0.027***	0.044***
	(0.002)	(0.002)
_cons	0.146***	0.598***
	(0.011)	(0.018)
N	3,416.000	3,416.000
R-squared	0.436	0.360
F-stat	404.11	7.43

Notes: Robust standard errors in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

With regard to Condition 2, this study provides indirect empirical support through a series of verifiable falsification tests, specifically including density continuity tests and covariate balance tests.

To rule out the possibility of individuals deliberately manipulating their reported age, this study first conducts a density continuity test on the driving variable (age). This paper employs the new local polynomial density estimation method developed by Cattaneo et al. [35] based on the aforementioned methods. It is intuitive, easy to implement, and entirely data-driven.

In practice, to avoid any potential impact of bandwidth selection on the test results, this study manually sets the optimal bandwidths asymmetrically for the left and right sides (5 for the left and 15 for the right). The test results show that the difference in density estimates at the cutoff is statistically highly insignificant (z = 0.653, p = 0.514). This result indicates that it cannot reject the null hypothesis. The distribution on both sides

of the cutoff is smooth and continuous, and there is no evidence of individuals precisely manipulating their reported ages. This provides a prerequisite for the validity of the fuzzy regression discontinuity design.

Based on the results of the density test, this study further conducted a covariate balance test [36] to rule out the possibility that other policy or human factors caused a sudden change at the cutoff.

This study used educational level, marital status, and residence status as dummy variables, applied the same kernel function and bandwidth selection methods as in the baseline regression, and conducted local polynomial regression at the age of 50. Specifically, educational level and marital status follow the same categorization used for the control variables described earlier. Residence status, however, is recategorized based on whether the individual has an agricultural HuKou. Table 2 reports the three types of estimators and their standard errors. The results show that none of the estimated coefficients for the three dummy variables reach statistical significance, and all estimates are small. This indicates that, near the cutoff, the distribution of the samples across the pre-determined characteristics is continuous, and there is no selection bias.

Based on the results of these two tests, this study found no direct evidence of a violation of the validity assumption, laying the groundwork for subsequent tests such as benchmark regression.

Table 2. Covariate balance test

	(1)	(2)	(3)
	Educational level	Marital status	Residential status
Conventional	0.195	-0.061	0.020
	(0.165)	(0.077)	(0.141)
Bias-corrected	0.413	-0.045	0.007
	(0.165)	(0.077)	(0.141)
Robust	0.413	-0.045	0.007
	(0.267)	(0.117)	(0.224)

Notes: Robust standard errors in parentheses.

9. Empirical results

To test the stability of the results, this paper reports the regression results under both settings: with and without control variables. Given that the inclusion of control variables effectively absorbs macroeconomic shocks across years and accounts for differences in individual sociodemographic characteristics, this paper adopts the models with control variables (Models (3) and (4) in Table 3) as the benchmark regression models. The results of subsequent robustness and falsification tests are compared with those of this model.

Models 1 and 2 of Table 3 present the regression results for the effects of retirement on different cognitive function modules in women. None of the estimates in the models includes covariates, and the bandwidth was selected using the Mean Squared Error (MSE)-optimized bandwidth selection method. For the mental intactness module, all three estimates are negative and none of the three results are significant. For the episodic memory module, all of the three estimates are also not significant.

Models 3 and 4 of Table 3 present the fuzzy RDD results for the effects of retirement on cognitive functions among women after controlling for covariates. This analysis also employs the MSE-based optimal bandwidth selection method. In the analysis of mental intactness module, the absolute values of the three estimates shows a slight increase, but none reached statistical significance. For episodic memory module, the coefficients are also negative and not significant.

In conclusion, regardless of whether control variables are included, the estimated coefficients for retirement on women's mental intactness and episodic memory do not reach statistically significant levels. These indicate that no significant net causal effect between retirement and women's cognitive function is observed around the statutory retirement age.

Table 3. Fuzzy Regression Discontinuity Design (RDD) estimates for women

	Without covariates		With covariates	
	(1) Mental Intactness	(2) Episodic Memory	(3) Mental Intactness	(4) Episodic Memory
Conventional	-0.255	-0.002	-0.394	-0.231
	(0.290)	(0.330)	(0.293)	(0.320)
Bias-corrected	-0.193	0.088	-0.407	-0.197
	(0.290)	(0.330)	(0.293)	(0.320)
Robust	-0.193	0.088	-0.407	-0.197
	(0.473)	(0.547)	(0.469)	(0.509)
Bandwidths & Sample size				
Main bandwidth (h)	3.540	3.575	3.358	3.242
Bias bandwidth (b)	4.981	4.817	4.864	4.832
Effective sample size (Left / Right)	500/711	475/682	500/711	475/682

Notes: Clustered standard errors in parentheses.

Table 3 only reports the estimated coefficients for the core explanatory variable. The complete regression results presents in Appendix 1. The direction of the coefficients is consistent with theoretical expectations. No findings were found to contradict theoretical expectations. This supports the validity of the baseline model.

10. Robustness test

To test the stability and reliability of the regression estimates, this paper conducts robustness tests by adjusting the bandwidth, kernel function, and polynomial order [20]. The control variable models present in Models (3) and (4) of Table 3 serve as the baseline.

In a fuzzy regression discontinuity design, the choice of bandwidth is critical, and regression results are often highly sensitive to changes in this parameter [36]. This paper selects the CER-based optimal bandwidth selection method to replace the original MSE-based bandwidth selection method. As shown in Models 1 and 2 of Table 4, after applying the optimal CER bandwidth, none of the estimates reaches the significance level. These indicate that the choice of bandwidth do not alter the baseline conclusions.

Second, this paper replaces the triangular kernel with a uniform kernel. When combined with MSE bandwidth, both triangular kernel function and uniform kernel function are optimal. The results show that the conventional and bias-corrected estimates for mental intactness are significant at the 10% and 5% levels, respectively. The direction of the estimated coefficients for episodic memory is opposite to that of the baseline model. But, none of the three estimates reaches statistical significance (see Models 3 and 4 of Table 4).

Third, this study employs the effect of the polynomial degree. In robustness test, this paper examines the results under a quadratic polynomial. As shown in Models 5 and 6 of Table 4, the direction of bias-corrected

estimates for mental intactness changed from negative to positive, and none of the estimates reaches the level of significance. The bias-corrected estimate for the episodic memory module is significant at the 1% level.

Based on the results of the robustness tests described above, when the bandwidth selection method, kernel function, and local polynomial order are changed, the absolute values and signs of the estimated coefficients for each model fluctuate to some extent. Yet most of the estimated results are not statistically significant and remain consistent with the baseline model. Isolated significant estimates under alternative specifications do not overturn the baseline null.

Table 4. Robustness test

	CER		Uniform Kernel		Quadratic	
	(1)	(2)	(3)	(4)	(5)	(6)
	Mental Intactness	Episodic Memory	Mental Intactness	Episodic Memory	Mental Intactness	Episodic Memory
Conventional	-0.253	-0.445	-0.518*	0.051	-0.053	-0.819
	(0.356)	(0.390)	(0.272)	(0.286)	(0.547)	(0.634)
Bias-corrected	-0.278	-0.389	-0.662**	0.284	0.438	-1.835***
	(0.356)	(0.390)	(0.272)	(0.286)	(0.547)	(0.634)
Robust	-0.278	-0.389	-0.662	0.284	0.438	-1.835
	(0.486)	(0.536)	(0.458)	(0.498)	(1.048)	(1.233)
Bandwidths & Sample size						
Main bandwidth (h)	2.292	2.217	3.338	3.292	3.650	3.260
Bias bandwidth (b)	4.864	4.832	4.797	4.614	4.971	4.509
Effective sample size (Left / Right)	348/543	331/522	500/711	475/682	500/711	475/682

Notes: Clustered standard errors in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table 5. Individual fixed effects estimates

	(1)	(2)
	Mental Intactness	Episodic Memory
retired	-0.102	-0.020
	(0.119)	(0.103)
_cons	-0.494	-0.174
	(0.655)	(0.469)
N	1,211.000	1,157.000
R-squared	0.049	0.087

Notes: Clustered standard errors in parentheses.

Furthermore, Table 5 reports robustness tests using a local linear regression model with individual fixed effects. The results show that the estimated coefficient for retirement's effect on women's psychological well-being is −0.102 (standard error = 0.119), and the estimated coefficient for retirement's effect on episodic memory is −0.020 (standard error = 0.103). Both coefficients are not statistically significant. These results are consistent with the findings of the baseline regression, indicating that the baseline results are not driven by overlooked individual-level fixed factors. This further enhances the robustness of the current study's conclusions.

11. Donut test

Although model validity tests confirmed that there was no artificial manipulation of age at the cutoff point, middle-aged and older adults in China tend to report ages ending in "0" (i.e., whole decades) [37]. Therefore, age stacking may occur. To exclude this possibility, this paper follows the method proposed by Barreca et al. [38] and applies donut regression. This paper excludes samples with a standardized age of 0, i.e., 50 years old. The test still uses the model with control variables included in Table 3 as the baseline model.

As shown in Models 1 and 2 of Table 6, the results of the donut test indicate that the absolute values of the coefficients for mental intactness increase, but their estimated directions remain the same, while the estimated directions for the episodic memory module are opposite to those of the baseline regression. In terms of significance, the estimates for mental intactness are significant at the 5%, 1%, and 5% levels, respectively. The estimates for the episodic memory module remain insignificant.

Table 6. Donut test and placebo test

	Delete those aged 50		48 Years Old		52 Years Old	
	(1)	(2)	(3)	(4)	(5)	(6)
	Mental Intactness	Episodic Memory	Mental Intactness	Episodic Memory	Mental Intactness	Episodic Memory
Conventional	-0.769**	0.174	-3.853	3.623	-2.201	0.928
	(0.322)	(0.295)	(4.886)	(3.588)	(1.665)	(1.157)
Bias-corrected	-0.840***	0.088	-12.250**	19.212***	-1.405	0.906
	(0.322)	(0.295)	(4.886)	(3.588)	(1.665)	(1.157)
Robust	-0.840**	0.088	-12.250	19.212**	-1.405	0.906
	(0.425)	(0.405)	(11.334)	(8.425)	(2.585)	(1.798)
Bandwidths & Sample size						
Main bandwidth (h)	4.033	4.035	3.358	3.358	3.242	3.242
Bias bandwidth (b)	7.990	8.069	4.864	4.864	4.832	4.832
Effective sample size (Left / Right)	638/684	608/656	388/732	367/699	556/677	533/650

Notes: Clustered standard errors in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Models 1 and 2 in Table A2 present the full regression results from the donut test. The results show that, after excluding the 50-year-old sample, the signs and significance levels of the coefficients for each control variable are largely consistent with those in the baseline Table A1 and are in line with theoretical expectations.

To examine the samples' characteristics at the cutoff, this paper presents a histogram of the age distribution (Figure A3). The data show that there is no significant concentration of age within the sample at age 50, thereby ruling out the possibility of age manipulation. However, the individuals at the age-50 are those who have just completed the retirement procedures. The cognitive changes among this group of recent retirees may not yet be fully manifested. Therefore, the donut test actually reflects differences between individuals who have not yet retired and those who have been retired for some time. Compared to the baseline's immediate net effect, the significant effect observed in the donut test is consistent with the delayed impact of retirement on mental intactness. At the time of retirement, positive and negative mechanisms may offset each other, while negative mechanisms may gradually accumulate as time elapses after retirement. However, because this study uses pooled cross-sectional data, it does not track data for the same respondents across different time periods. Therefore, the estimates from the donut test cannot be interpreted as strict lag effects. They provide only suggestive evidence consistent with a potential delayed manifestation. Therefore, the findings of the donut test complement rather than contradict the baseline findings, but they still require further validation using longer-term individual-level panel data.

12. Placebo test

To further validate the specificity of the benchmark results and exclude the possibility that the conclusions are due to age-related confounding factors or coincidental model specifications, this study follows existing methods to conduct placebo tests by setting artificial cutoff values, different from the actual cutoff value, in intervals below and above the true cutoff value [36]. Consistent with the previous analysis, this section continues to use the specifications of Models (3) and (4) in Table 3 as the benchmark. The placebo test moves the cutoff to a manufactured position to verify whether the baseline results occur only at the true age-50 cutoff. In this placebo test, the cutoffs are set at ages 48 and 52.

Because there are a large number of mass points near the two pseudo-cutoffs in the running variables, the MSE method for selecting the optimal bandwidth cannot be implemented properly. This study uses the method proposed by Noack and Rothe to hand-set the bandwidth in line with the benchmark regression[3]. The test results are shown in Table 6 from Model 3 to Model 6.

As shown in Models 3 and 4 of Table 6, when 48 years of age is set as the pseudo-cutoff, the absolute values of the estimated coefficients for mental intactness increase, and the standard errors expand. The bias-corrected estimate is significant at the 5% level. The absolute values and standard errors of the estimated coefficients for the episodic memory module also increase, with the direction of the estimates opposite to that of the baseline model. The bias-corrected estimate is significant at the 1% level, the robust estimate is significant at the 5% level, while the conventional estimate does not reach the level of significance.

As shown in Models 5 and 6 of Table 6, the absolute values and standard errors of the estimated coefficients for the two cognitive function modules at the pseudo-cutoff point of age 52 are amplified, but the extent of this amplification is smaller than when the pseudo-cutoff point is set at age 48. The direction of

[3] Noack, C., & Rothe, C. (2023). Donut regression discontinuity designs (arXiv:2308.14464). arXiv. https://arxiv.org/abs/2308.14464. This paper is an arXiv preprint, provided here as supplementary reference for bandwidth setting methods.

change is the same as when the pseudo-cutoff point is set at age 48. None of the estimates reaches the level of significance.

The significant positive effect observed in the 48-year-old placebo test is consistent with the "on-the-job retirement" effect described in the literature [9]. This leads to a temporary improvement in episodic memory. The disappearance of the pseudo-cutoff effect at age 52 further confirms the central role of the age 50 as an institutional shock.

Models 3 through 6 of Table A2 present the full regression results from the placebo test. With a pseudo-cutoff set at age 48 (Models 3 and 4), the performance of the control variables is essentially consistent with that shown in Table A1 of the baseline study. Under the pseudo-cutoff at age 52 (Models 5 and 6), the estimated coefficients and their signs for some control variables change, and the standard errors expand significantly. This fluctuation may come from the fact that age 52 is not the actual statutory retirement age. There is no institutional change at this point, and the local regression estimates primarily reflect random noise in the data.

13. Discussion

This study found no significant net effect of retirement on the cognitive function of urban women in China affected by the age 50 statutory retirement policy.These results remain consistent across different models, bandwidth choices, and robustness tests. This study provides reliable causal estimates through a fuzzy RDD design. The following discussion builds upon existing literature and theoretical frameworks to provide mechanistic and inferential explanations for the non-significant findings of this study.

The relationship between retirement and cognitive function has been a subject of debate in the existing literature. This study provides further empirical evidence regarding this inconsistency.

The discrepancy and connection between the null results of this study and the findings of other studies, which are also based on Chinese microdata that examine the impact of retirement on cognitive function can be explained from the following two perspectives. First, the differences from Chen [18] and Xiong et al. [19] stem from the different objects of identification. These two studies examine the spillover effects of a spouse's retirement, which are highly dependent on patterns of interaction and division of labor within the household and are susceptible to confounding by household heterogeneity, leading to fluctuating conclusions. This paper, however, examines the direct effects of the statutory retirement age on women themselves. Second, this paper further confirms that the null result reported by Wang et al. [20] remains robust among urban women subject to the age-50 retirement policy in China.

However, the non-significant results of this study do not indicate that retirement has no effect on cognition. The impact of retirement itself is two-sided. Its effects on cognitive function may be offset by multiple opposing forces. At the same time, cognitive decline is shaped by a combination of many factors, which may dilute the marginal effect of retirement.

To provide a preliminary description of the pathways influencing emotional health, this study examined the total CESD-10 scores and feelings of loneliness among individuals around the age of 50 (Figure 3). The figure shows that CESD-10 scores rise after age 50, indicating an overall increase in depressive symptoms. Feelings of loneliness, on the other hand, tend to decline after this age. This pattern suggests that retirement may influence individuals' emotional well-being in two opposing ways.

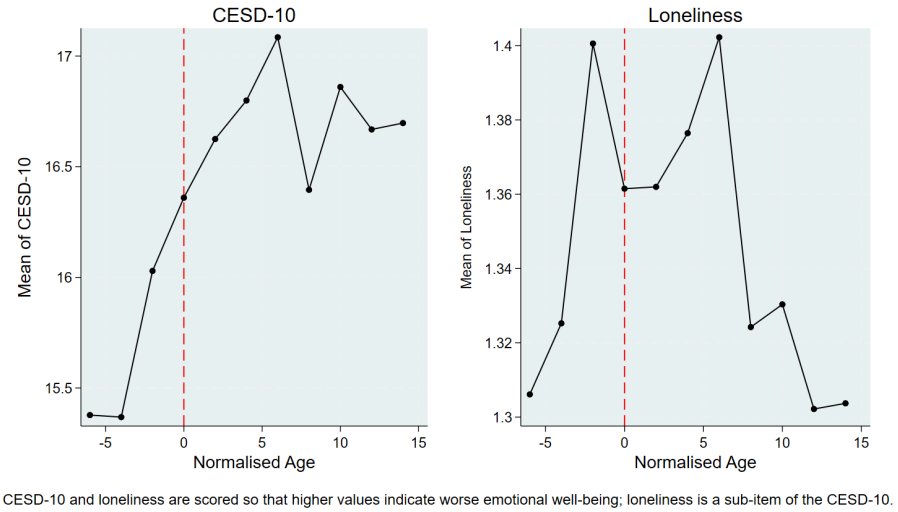


Figure 3. Emotional well-being around the age-50 cutoff

On the one hand, retirement frees individuals from the high stress of the workplace and may improve mental health. In addition to a sharp reduction in work-related stress, many older adults also have ample time to adopt healthier lifestyles. Atalay et al. [39] also found that retirement significantly increases social capital and strengthens older adults' social connections and community engagement. These behavioral changes further contribute to emotional stability and overall health, and reduce feelings of loneliness. These improvements in mood may indirectly create favorable conditions for maintaining cognitive function.

Conversely, retirement may also worsen one's emotional state and lead to a decline in cognitive function. The increase in CESD-10 scores also suggests that depressive symptoms may increase after retirement. Stopping work brings changes in social identity, source of income, which may lead to a loss of a sense of accomplishment and an increase in symptoms of depression. Long-term depression constitutes a risk factor for cognitive decline. Therefore, retirement may also have an indirect negative effect on cognition through the pathway of deteriorating mood.

Retirement frees women from formal, paid work, but they often immediately transition into unpaid family caregiving. Therefore, this retired but active state objectively provides an alternative source of cognitive stimulation. A study by Zhou and Hong [40] found that non-intensive intergenerational caregiving is associated with a lower risk of dementia. Reducing feelings of loneliness is one of the reasons for this.

However, this cognitive substitution effect has its costs. For retired women, becoming family caregivers is not always a voluntary choice. Rather, it is an obligation under the dual influence of traditional gender norms and notions of successful aging. The concept of successful aging encourages older adults to remain active and socially engaged. However, this ideal is gendered within the family. Families translate it to expect retired women to maintain family relationships and provide care and daily support [41]. The lack of external caregiving resources further reinforces this gendered inequality [41]. Empirical research has confirmed the psychological burden of this dilemma. Spouses who provide care for Activities of Daily Living (ADL) experience a faster decline in cognitive function, particularly among women [42]. When caregiving shifts from a choice to a socially imposed obligation, the positive influences on cognitive may be masked by external pressures. Once caregiving intensity exceeds the threshold that an individual can tolerate, this cognitive substitution effect may be offset by psychological burden and may further evolve into cognitive impairment.

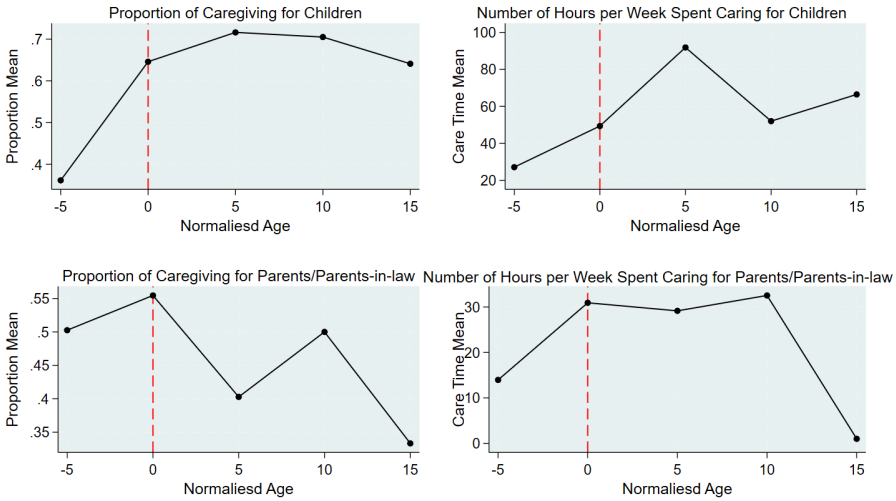


Figure 4. Family caregiving situations around the age-50 cutoff

This paper also provides a descriptive analysis of the family caregiving module (Figure 4). The descriptive analysis of family caregiving is limited to the 2011, 2013, and 2015 waves. This is because the 2020 CHARLS data lack comparable caregiving variables, and the caregiving variables for parents in the 2018 data were recorded differently from those in earlier waves. Using the same sample definition as in the main analysis, I examined two dimensions of family caregiving: caregiving for children and caregiving for parents or parents in-laws. For each dimension, this study plotted the group-level means of care participation rates and weekly care hours by standardized age. This is a descriptive analysis and cannot represent the results of a causal mechanism analysis.

Figure 4 shows that the proportion of women providing childcare increases after age 50, and the number of hours spent caring for children each week also rises. In contrast, the proportion of women providing care for their parents or in-laws tends to decline after age 50. This suggests that retirement does lead women to take on more childcare for their grandchildren. While this provides another form of cognitive stimulation, it may also bring burdens that offset these benefits. However, because the caregiving variable is self-reported and was measured only in the early survey waves, there may be endogeneity issues here. Therefore, these results should be viewed as descriptive evidence rather than a formal test of the mechanisms.

In general, descriptive evidences on emotional well-being and family caregiving suggest that these two paths are closely linked to the lives of urban Chinese women after retirement, particularly for blue-collar or non-managerial women affected by the age-50 retirement policy. The phenomenon of rising CESD-10 scores accompanied by a simultaneous decline in feelings of loneliness also reflects the dual nature of the shift in women's social roles after retirement. On the one hand, as shown in Figure 4, the proportion of women participating in childcare rises after age 50, providing them with another avenue for social interaction and a sense of purpose in life, which helps alleviate the social isolation that may result from losing connections in the workplace; this is consistent with the observed decline in feelings of loneliness. However, retired women are often expected to undertake heavier family caregiving responsibilities. The resulting mental strain, combined with the sudden loss of personal income and social status, creates a persistent psychological burden. This may exacerbate symptoms of depression. As a result, the positive emotional effects of reducing loneliness are offset by the negative emotional costs because of caregiving obligations. Thereby, their overall effect on cognition may be weakened.

Except retirement, cognitive function is influenced by other factors. Age is the strongest factor contributing to cognitive decline. Even ignoring other environmental or behavioral factors, the aging process may lead to a natural decline in cognitive function. This study employs a fuzzy regression discontinuity design to control for this decline. However, there are still a wide range of factors which are closely linked to cognitive function, such as socioeconomic status [43], educational attainment [43], marital status [44], social engagement [45], physical health [6, 7], sleep quality, dietary habits [46], and even psychological factors such as emotional exhaustion and rumination [47, 48].

In summary, the failure to find a significant effect of retirement on women's cognitive function is not because retirement is unrelated to cognitive function, but rather because its impact is embedded within a multidimensional causal network.

In addition, the non-significant results of this study may also be related to the mixed composition of the sample at the age-50 cutoff. At the age of 50, some female white-collar workers subject to the age-55 retirement policy remained in the control group and continued working. This group of women typically has a higher socioeconomic status and greater cognitive reserves, and their continued participation in the workforce may yield positive cognitive benefits. This may dilute or mask the true effect of blue-collar women's retirement, thereby causing the net effect to be statistically insignificant.

However, the above discussion is a theoretical interpretation based on existing literature. This study has controlled for the balance of observable covariates at the cutoff. It remains to be verified in future studies whether other factors not included in the analysis also changed simultaneously at the cutoff.

This study has the following limitations, which warrant further exploration in future research. First, retirement is narrowly defined. This study defines retirement status solely based on whether retirement procedures have been completed, and does not distinguish among dimensions such as the type of retirement (voluntary or involuntary), the duration of retirement, or whether the individual has returned to work after retirement.

Second, there are limitations in mechanism testing and dynamic validation. Although this paper uses the CESD-10 and the loneliness curve to provide descriptive evidence of emotional pathways, it does not conduct a formal test of the underlying mechanisms. Furthermore, since these variables are all self-reported data, there may be issues of endogeneity. In addition, the CHARLS 2020 wave lacks comparable caregiving variables, and the 2018 wave's parent-care variables are not comparable to those in earlier waves. Therefore, the descriptive analysis of family caregiving is limited to the 2011, 2013, and 2015 waves. As a result, it is impossible to accurately distinguish between the relative contributions of the improvement path and the deterioration path. Future research should employ measurement indicators that are consistent across waves and utilize formal mediation models or dynamic panel models for analysis. In addition, the donut regression results suggest that the negative impact of retirement on mental intactness may manifest with a delay. However, given the design of this study, this result cannot be interpreted as a strict lagged effect and should be regarded as suggestive evidence. Future research could use long-term individual panel data to analyze the changes of cognitive function after retirement.

Third, there are limitations in variable usage and extension to other groups. This study is unable to separately identify the groups of women subject to the age 50 and age 55 retirement policies. Although this study has taken measures to solve this issue, such as defining the LATE, controlling for covariates, and conducting robustness tests, the estimation results may still retain residual heterogeneity. Therefore, the findings of this study apply only to blue-collar urban women subject to the age 50 retirement system and cannot be generalized to women who are not enrolled in formal pension plans or affected by the age 55 retirement policy. If future research can obtain detailed data on retirement year records or occupational

categories when retirement, it would be possible to further disentangle the heterogeneity in cognitive effects and estimate it across different retirement systems and different types of populations.

14. Conclusion and policy recommendation

The core finding of this study is that, among urban women enrolled in the specified pension schemes and subject to the age-50 statutory retirement age, no robust evidence is found—regardless of control variables—that retirement has a significant immediate net effect on either mental intactness or episodic memory. Because the age-50 cutoff identifies a LATE for women covered by the age-50 retirement policy, mainly blue-collar or non-managerial urban female employees, these conclusions should not be generalized to all urban women in China or to women subject to the age-55 retirement policy.

Future research could utilize long-term individual follow-up data to investigate the dynamic changes in cognitive function after retirement.

Based on the mechanistic analysis in the discussion section, retirement and cognitive function are not unrelated. However, the independent effect of retirement on cognitive function is difficult to detect statistically. Therefore, this study argues that policymakers should address the multidimensional needs of retired women subject to the age-50 statutory retirement age, particularly blue-collar or non-managerial urban female employees, by providing alternative opportunities for cognitive stimulation and social engagement, while also addressing gender inequalities in the burden of family caregiving. First, policymakers should improve community care services to alleviate the burden of family caregiving on women. At the same time, they should encourage family members to share caregiving responsibilities. Second, they should develop channels such as senior universities and community volunteer programs to provide them with a wider range of social participation options.

References

[1] White, L., Arp, P. C., & Coe, N. B. (2025). Retirement and cognitive function in later life (Wharton Pension Research Council Working Paper No. 11). *Wharton Pension Research Council*. https://doi.org/10.2139/ssrn.5334947

[2] Baker, J., MacConnachie, L., Hood, M., Derby, C. A., Appelhans, B. M., Reeves, A., & Karvonen-Gutierrez, C. A. (2024). Physical function and cognition in older adult women: The study of women's health across the nation (SWAN). *Alzheimer's & Dementia, 20*(Suppl. 3), e086153. https://doi.org/10.1002/alz.086153

[3] Zhao, X., Tang, H., Duan, X., Shen, S., & Wang, J. (2025). Relationship between anthropometric indicators and cognitive function in older women. *Health Care for Women International, 46*(12), 1409–1423. https://doi.org/10.1080/07399332.2024.2413448

[4] Sheets, K. M., Kats, A. M., Fink, H. A., Langsetmo, L., Yaffe, K., & Ensrud, K. E. (2025). Life-space mobility and cognition in community-dwelling late-life women: A cross-sectional analysis. *Journal of the American Geriatrics Society, 73*(1), 206–213. https://doi.org/10.1111/jgs.19190

[5] Han, Y., Huang, L., Zhou, Y., Zhu, Y., Huang, S., Huang, X., Li, D., Yan, J., & Su, W. (2025). Sex-specific pathogenesis of Alzheimer's disease in elderly women: Mechanistic insights and therapeutic horizons. *Aging Research, 3*(1), 9340056. https://doi.org/10.26599/AGR.2025.9340056

[6] Wang, S.-M., Jeong, C., Um, Y. H., Kang, D. W., Kim, S., Lee, S., Lee, C. U., Aizenstein, H., & Lim, H. K. (2025). Follicle-stimulating hormone linked to cognitive decline and amyloid burden in postmenopausal women. *Frontiers in Aging Neuroscience, 17*, 1697255. https://doi.org/10.3389/fnagi.2025.1697255

[7] Gallay, C., Soldevila-Domenech, N., López-Martos, D., Sanchez-Benávides, G., Huguet, J., Suárez-Calvet, M., Milà-Alomà, M., Grau-Rivera, O., & the ALFA Study Group. (2026). Effect of parity and β-amyloid on cognition and hippocampal volume in postmenopausal women. *Neurology, 107*(1), e218153. https://doi.org/10.1212/WNL.0000000000218153

[8] Dong, J., Su, C., Zhang, R., Luo, X., Huang, R., Sun, B., Li, J., & Tang, M. (2025). A follow-up study on the trajectory and sex differences of different cognitive dimensions in the normal aging elderly in Guangzhou. *Alpha Psychiatry, 26*(3), 44186. https://doi.org/10.31083/AP44186

[9] Rohwedder, S., & Willis, R. J. (2010). Mental retirement. *Journal of Economic Perspectives, 24*(1), 119–138. https://doi.org/10.1257/jep.24.1.119

[10] Nie, Y., Richards, M., Kubinova, R., Titarenko, A., Malyutina, S., Kozela, M., Pajak, A., Bobak, M., & Ruiz, M. (2021). Social networks and cognitive function in older adults: Findings from the HAPIEE study. *BMC Geriatrics, 21*(1), 570. https://doi.org/10.1186/s12877-021-02531-0

[11] Sato, K., Noguchi, H., & Inoue, K. (2025). Heterogeneity in the association between retirement and cognitive function: A machine learning analysis across 19 countries. *International Journal of Epidemiology, 54*(6), dyaf201. https://doi.org/10.1093/ije/dyaf201

[12] Coleman, M. E., Roessler, M. E. H., Peng, S., Roth, A. R., Risacher, S. L., Saykin, A. J., Apostolova, L. G., & Perry, B. L. (2023). Social enrichment on the job: Complex work with people improves episodic memory, promotes brain reserve, and reduces the risk of dementia. *Alzheimer's & Dementia, 19*(6), 2655–2665. https://doi.org/10.1002/alz.13035

[13] Greenberg, K., & Burgard, S. (2021). Cumulative employment intensity and complexity across the life course and cognitive function in later life among European women and men. *Annals of Epidemiology, 58*, 83–93. https://doi.org/10.1016/j.annepidem.2021.01.006

[14] Xue, B., Pai, M., & Luo, M. (2022). Working beyond SPA and the trajectories of cognitive and mental health of UK pensioners: Do gender, choice, and occupational status matter? *European Journal of Ageing, 19*(3), 423–436. https://doi.org/10.1007/s10433-021-00644-4

[15] Schwartz, G. L., & Glymour, M. M. (2023). Bridging the divide: Tackling tensions between life-course epidemiology and causal inference. *Annual Review of Developmental Psychology, 5*, 355–374. https://doi.org/10.1146/annurev-devpsych-120221-033623

[16] Hale, J. M., Bijlsma, M. J., & Lorenti, A. (2021). Does postponing retirement affect cognitive function? A counterfactual experiment to disentangle life course risk factors. *SSM - Population Health, 15*, 100855. https://doi.org/10.1016/j.ssmph.2021.100855

[17] Shin, S. H., & Fan, J. X. (2025). Gender-specific impacts of employment on mental and cognitive health in older adults. *Aging & Mental Health, 29*(10), 1821–1830. https://doi.org/10.1080/13607863.2025.2502797

[18] Chen, X. (2022). The impact of spousal and own retirement on health: Evidence from urban China. *World Development, 159*, 106025. https://doi.org/10.1016/j.worlddev.2022.106025

[19] Xiong, X., Li, R., & Yang, H. (2021). How the spouse's retirement affects the cognitive health of individuals in China: A fresh evidence from the perspective of social interaction. *Frontiers in Public Health, 9*, 796775. https://doi.org/10.3389/fpubh.2021.796775

[20] Wang, T., Liu, H., Zhou, X., & Wang, C. (2024). The effect of retirement on physical and mental health in China: A nonparametric fuzzy regression discontinuity study. *BMC Public Health, 24*(1), 1184. https://doi.org/10.1186/s12889-024-18649-w

[21] Lee, H., Lim, J. A., & Nam, H. K. (2022). Effect of a digital literacy program on older adults' digital social behavior: A quasi-experimental study. *International Journal of Environmental Research and Public Health, 19*(19), 12404. https://doi.org/10.3390/ijerph191912404

[22] Kim, J. H., Muniz-Terrera, G., & Leist, A. K. (2023). Does (re-)entering the labour market at advanced ages protect against cognitive decline? A matching difference-in-differences approach. *Journal of Epidemiology and Community Health, 77*(10), 663–669. https://doi.org/10.1136/jech-2022-220197

[23] Giles, J., Wang, D., & Cai, W. (2012). The labor supply and retirement behavior of China's older workers and elderly in comparative perspective. In J. P. Smith & M. Majmundar (Eds.), *Aging in Asia: Findings from new and emerging data initiatives* (pp. 116–147). National Academies Press.

[24] Zhu, H., & Walker, A. (2018). Pension system reform in China: Who gets what pensions? *Social Policy & Administration, 52*(7), 1410–1424. https://doi.org/10.1111/spol.12368

[25] Zhao, Y., Hu, Y., Smith, J. P., Strauss, J., & Yang, G. (2014). Cohort profile: The China Health and Retirement Longitudinal Study (CHARLS). *International Journal of Epidemiology, 43*(1), 61–68. https://doi.org/10.1093/ije/dys203

[26] Lei, X., Hu, Y., McArdle, J. J., Smith, J. P., & Zhao, Y. (2012). Gender differences in cognition among older adults in China. *Journal of Human Resources, 47*(4), 951–971. https://doi.org/10.3368/jhr.47.4.951

[27] Li, H., Li, C., Wang, A., Qi, Y., Feng, W., Hou, C., Tao, L., Liu, X., Li, X., Wang, W., Zheng, D., & Guo, X. (2020). Associations between social and intellectual activities with cognitive trajectories in Chinese middle-aged and older adults: A nationally representative cohort study. *Alzheimer's Research & Therapy, 12*(1), 115. https://doi.org/10.1186/s13195-020-00691-6

[28] Wang, J., Cui, Q., Xu, X., & Yang, G. (2025). Bidirectional association between grip strength and cognitive function in Chinese older adults: A nationwide cohort study. *BMC Public Health, 25*(1), 1880. https://doi.org/10.1186/s12889-025-23079-3

[29] Meng, Q., Wang, H., Strauss, J., Langa, K. M., Chen, X., Wang, M., Qu, Q., Chen, W., Kuang, W., Zhang, N., Li, T., Wang, Y., & Zhao, Y. (2019). Validation of neuropsychological tests for the China Health and Retirement Longitudinal Study Harmonized Cognitive Assessment Protocol. *International Psychogeriatrics, 31*(12), 1709–1719. https://doi.org/10.1017/S1041610219000693

[30] Allen, M. P. (1997). *Understanding regression analysis*. Springer. https://doi.org/10.1007/b102242

[31] Lee, D. S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature, 48*(2), 281–355. https://doi.org/10.1257/jel.48.2.281

[32] Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression-discontinuity design. *Econometrica, 69*(1), 201–209. https://doi.org/10.1111/1468-0262.00183

[33] Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*(3), 557–586. https://doi.org/10.2307/2171753

[34] Stock, J. H., & Yogo, M. (2005). Testing for weak instruments in linear IV regression. In D. W. K. Andrews & J. H. Stock (Eds.), *Identification and inference for econometric models: Essays in honor of Thomas Rothenberg* (pp. 80–108). Cambridge University Press. https://doi.org/10.1017/CBO9780511614491.006

[35] Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association, 115*(531), 1449–1455. https://doi.org/10.1080/01621459.2019.1635480

[36] Cattaneo, M. D., & Titiunik, R. (2022). Regression discontinuity designs. *Annual Review of Economics, 14*, 821–851. https://doi.org/10.1146/annurev-economics-051520-021409

[37] Gu, D., & Feng, Q. (2019). Use of the average age ratio method in analyzing age heaping in censuses: The case of China. *International Journal of Population Studies, 5*(1), 13–26. https://doi.org/10.18063/ijps.v5i1.979

[38] Barreca, A. I., Guldi, M., Lindo, J. M., & Waddell, G. R. (2011). Saving babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics, 126*(4), 2117–2123. https://doi.org/10.1093/qje/qjr042

[39] Atalay, K., Staneva, A., & Zhu, R. (2026). Integration or isolation? The impact of retirement on social capital. *Journal of Human Resources*. Advance online publication. https://doi.org/10.3368/jhr.0625-14357R2

[40] Zhou, Z., & Hong, X. (2025). Caring for grandchildren and dementia among older adults in China. *JAMA Network Open, 8*(7), e2519622. https://doi.org/10.1001/jamanetworkopen.2025.19622

[41] Ti, L. (2026). She never really retired: Successful aging, emotional labor, and gendered social reproduction in China. *Women's Studies International Forum, 119*, 103395. https://doi.org/10.1016/j.wsif.2025.103200

[42] Guo, Y., Zhang, Z., & Jiang, Q. (2025). Spousal caregiving types and cognitive trajectories among middle-aged and older adults in China. *Research on Aging, 47*(5–6), 308–320. https://doi.org/10.1177/01640275251317544

[43] Brown, K. E. (2025). Cognitive aging through a life course lens: The roles of childhood socioeconomic status, adult education, and aging self-perceptions in older women [Doctoral dissertation, Idaho State University]. *Idaho State University Electronic Dissertations and Theses*.

[44] Nerobkova, N., Park, Y. S., Shin, J., & Park, E. C. (2022). Marital transition and cognitive function among older adults: The Korean Longitudinal Study of Aging (2006–2020). *BMC Geriatrics, 22*(1), 1003. https://doi.org/10.1186/s12877-022-03697-x

[45] Tomioka, K., Kurumatani, N., & Hosoi, H. (2018). Social participation and cognitive decline among community-dwelling older adults: A community-based longitudinal study. *The Journals of Gerontology: Series B, 73*(5), 799–806. https://doi.org/10.1093/geronb/gbw059

[46] Berendsen, A. M., Kang, J. H., Feskens, E. J. M., de Groot, C. P. G. M., Grodstein, F., & van de Rest, O. (2018). Association of long-term adherence to the MIND diet with cognitive function and cognitive decline in American women. *The Journal of Nutrition, Health & Aging, 22*(2), 222–229. https://doi.org/10.1007/s12603-017-0909-0

[47] Baker, K., Taylor, W. D., & Szymkowicz, S. M. (2026). Brooding rumination predicts lower cognitive functioning in late-life depression: Multivariate multiple regression analyses. *Journal of Clinical and Experimental Neuropsychology, 48*(2), 142–155. https://doi.org/10.1080/13803395.2025.2520356

[48] Ye, N., Peng, L., Deng, B., Hu, H., Wang, Y., Zheng, T., Ai, Y., Liu, X., Zhou, S., & Li, Y. (2025). Repetitive negative thinking is associated with cognitive function decline in older adults: A cross-sectional study. *BMC Psychiatry, 25*(1), 562. https://doi.org/10.1186/s12888-025-06815-2

Appendix

Table A1. Complete coefficient report of benchmark RDD control variables

	(1)	(2)
	Mental Intactness	Episodic Memory
Non-Agricultural HuKou	0.319**	0.372***
	(0.132)	(0.138)
Unified Residence HuKou	0.198	0.158
	(0.184)	(0.194)
Do not Have HuKou	0.995**	0.787
	(0.441)	(0.499)
High_edu	0.448***	0.560***
	(0.074)	(0.078)
Having a Spouse	0.371**	0.163
	(0.173)	(0.163)
Wave-2013	0.377**	0.467***
	(0.163)	(0.164)
Wave-2015	0.414***	0.540***
	(0.148)	(0.148)
Wave-2018	0.123	0.959***
	(0.163)	(0.158)
Wave-2020	0.366**	1.106***
	(0.159)	(0.153)
N	1,211.000	1,157.000
R^2	0.098	0.164

Notes: Clustered standard errors in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

Table A2. Complete coefficient report for donut test and placebo test

	Delete those aged 50		48 Years Old		52 Years Old	
	(1) Mental Intactness	(2) Episodic Memory	(3) Mental Intactness	(4) Episodic Memory	(5) Mental Intactness	(6) Episodic Memory
Non-Agricultural HuKou	0.398***	0.196*	0.425***	0.161	-0.885	0.793
	(0.128)	(0.110)	(0.113)	(0.122)	(2.014)	(1.108)

Table A2. Continued

Unified Residence HuKou	0.318	0.013	0.293*	-0.094	-1.352	0.913
	(0.199)	(0.168)	(0.173)	(0.191)	(2.803)	(1.530)
Do not Have HuKou	1.394***	-0.022	1.398***	-0.364	0.000	0.000
	(0.387)	(0.313)	(0.530)	(0.565)	(.)	(.)
High_edu	0.433***	0.580***	0.409***	0.565***	0.895	0.291
	(0.080)	(0.071)	(0.074)	(0.080)	(0.739)	(0.416)
Having a Spouse	0.299	0.173	0.274*	0.184	0.131	0.193
	(0.210)	(0.122)	(0.156)	(0.152)	(0.276)	(0.190)
Wave-2013	0.392***	0.212*	0.383***	0.172	-2.246	1.377
	(0.140)	(0.128)	(0.136)	(0.139)	(4.523)	(2.407)
Wave-2015	0.415***	0.375***	0.341**	0.212	-1.747	1.338
	(0.133)	(0.123)	(0.134)	(0.145)	(3.754)	(2.070)
Wave-2018	0.118	0.793***	0.158	0.661***	-1.831	1.749
	(0.142)	(0.128)	(0.130)	(0.148)	(3.425)	(1.864)
Wave-2020	0.295**	0.929***	0.316**	0.932***	-1.741	1.754
	(0.139)	(0.126)	(0.134)	(0.148)	(3.620)	(1.946)
N	1,322.000	1,264.000	1,120.000	1,066.000	1,233.000	1,183.000
R^2	0.042	0.186	0.060	0.163	.	.

Note: R^2 is unavailable for the 52-year placebo specification due to computational constraints in the local weighted instrumental variable estimation. Do not Have HuKou omitted because of collinearity. Clustered standard errors in parentheses. ***, ** and * indicate statistical significance at the 1%, 5% and 10% levels respectively.

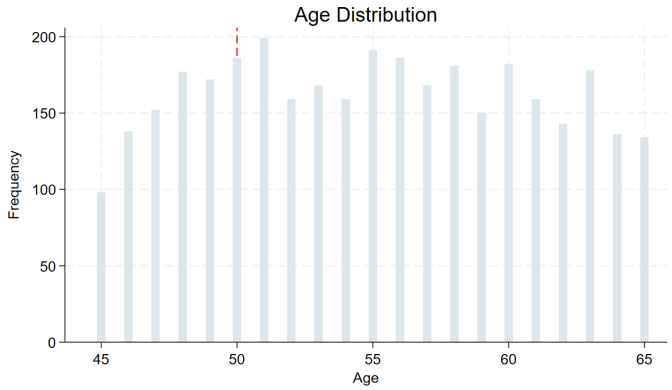


Figure A1. Age distribution of the sample