Depression Risk Prediction Based on IV-Probit and Transformer for Community Management

Yonghua Yang[1,a], Yarui Zhang[2,b], Tianyong Chen[3,c *]

{1045998686@qq.com[a], zhangyr@psych.ac.cn[b], * Corresponding author: chenty@psych.ac.cn[c] }

Institute of Psychology, Chinese Academy of Sciences, Department of Psychology, University of Chinese Academy of Sciences, Beijing, China[1]
Institute of Psychology, Chinese Academy of Sciences, Department of Psychology, University of Chinese Academy of Sciences, Beijing, China, Aba Teachers University, Sichuan, China[2]
Institute of Psychology, Chinese Academy of Sciences, Department of Psychology, University of Chinese Academy of Sciences, Beijing China[3]

Abstract. In modern society, community management has gained more attention on children, such as carrying out the three-child policy. On the contrary, the elderly seems be ignored, so this paper researches the correlation between the quantity of children and the well-being of the elderly. Employing data from the China Health and Retirement Longitudinal Study (CHARLS), this research employs an IV-Probit model to elucidate the correlation between the number of offspring and the aged mental well-being, specifically depression symptoms. The result suggests that (1) Chinese people have fertility preference for mix-sex; (2) and parents with more children had a remarkable negative impact on depression rate when endogeneity was not taken into account, but the offspring quantity had a non-significant protective effect on parental mental health when estimated using instrumental variables. In addition, with the aging of society, the mental health problems of the elderly have attracted much attention. This article uses the Transformer model to more accurately capture the diverse factors of depression risk in the elderly and make effective predictions through the advantages of powerful long-range dependency modeling, self-attention mechanism, and multi-head self-attention.

Keywords: community management, offspring quantity, gender structure of children, mental health of elderly

1 Introduction

The trend of population aging in China is characterized by a large base and rapid pace, but the construction of the corresponding old-age security system is clearly lagging behind and adverse to community management. Family support, especially children supporting will continue to exist and is the core of the old-age support which means that family old-age care is still the mainstay of our country's old-age care model[1], and the crucial role played by children in elderly care underscores their significance for improving parental physical and mental health [2]. However, the elderly people of the baby boom generation had more children, but successive generations have fewer children, raising concerns about the support available to new cohorts of elderly people, which may have an adverse effect on their mental health [3].

PMIS 2024, March 15-17, Changsha, People's Republic of China

DOI 10.4108/eai.15-3-2024.2346423

As further detailed consideration given later, adult children serve crucial resources for the aged people in China where there are potent familial norms and inadequate welfare systems. With multiple offspring, Chinese old people rarely face a disparity between actual and anticipated aid from their social networks. So, this investigation intends to illuminate whether increased birth numbers ameliorate late-life mental health risks within China's population.

Previously, studies investigating fertility - mental health interplay typically employed descriptive methodologies [4]. Careful explanation is warranted for prior study outcomes due to the socially influenced fertility and mental well-being. Case in point, unmeasured socioeconomic status and personal physique can potentially influence both reproductive activities and well-being, leading to biased results of prior studies. Traditional methods fail to consider underlying cross-cause interactions. This study drew from prior studies and employed the gender composition of the first two offspring as an instrumental variable, enable us to quantify the causal effect of high fertility rates on mental well-being.

In addition, with the arrival of an aging society, depression risk prediction in the elderly has become a research area that has attracted much attention. Traditional methods often employ statistics-based models, however, these methods have limitations in capturing complex long-range dependencies and diverse factors. To address this challenge, this paper proposes a Transformer-based depression risk prediction method for the elderly. Different from traditional methods, the Transformer model performs excellently in modeling long-range dependencies and sensitivity through its powerful self-attention mechanism and multi-head self-attention. This method more accurately captures the diverse factors of depression risk in older adults, providing a new and reliable tool for improving prediction accuracy and personalized intervention. Through the method proposed in this article, we can hope to more comprehensively understand and deal with mental health problems in older adults.

2 Theoretical Background and Hypothesis

Inquiry into the linkage between childbearing and mental health commenced primarily by public health scholars, psychologists, and sociologists. Numerous scholarly works dissected the prolonged impact of childbirth on mental well-being [5][6]. Sociologists emphasized the importance of children's social networks for aged parents [7]. Children can provide social aid along with routine assistance to their parents, so more children may be able to prevent feelings of loneliness in old age. Not only that, but parenthood provides individuals with a sense of fulfillment and also a sense of intent and significance in their lives, leading to enhanced welfare and improved mental well-being [8]. Yumeng Wang (2016), using the data from China Health and Retirement Longitudinal Study(CHARLS), focusing on the elderly’s life satisfaction, found a profound connection between the offspring quantity and increased life satisfaction, and childless individuals experienced notably decreased life satisfaction compared to parents [9]. Shenglong Liu (2020), by means of the data of China Family Panel Studies (CFPS), focusing on the old parents’ mental status in rural areas, found that the physiological state of parents who had two or more children were better than those who had only one child [10].

Some scholars have also provided micro- and macro-socio-psychological theories to explain why parenting is not connected with improved psychological state [11] [12]. Psychosocially

oriented mental health researchers have debated that the parental emotional gains could be offset [13], or outpaced by the emotional burdens [14], overshadowed by the numerous childcare stresses. These obligations, particularly during early childrearing, may undermine mental well-being [15]. Parenting may limit their ability to engage in other roles and incur sizable opportunity costs in the meantime. This cumulative stress could impose potential damage to mental health in later life, particular among large families or financial stressed households [16]. For mothers, children are even more of a source of stress, financial costs, and physical pain, so they are more likely to suffer from mental illness. Thus, there were some researches supporting different view from those mentioned above. Qiang Li (2021), also using data from CHARLS, focusing on the self rated score of happiness of old parents, tested the relationship and found that the effect was negatively related, and even was positively related with old parents' depression [17]. Dewei Geng (2013) using the data from a survey in urban area of Shanghai, Wuhan and Xian, focusing the self-rated physical condition of parents, tested the relationship and also found the effect was negatively related [18].

In the field of research on the effect of the number of children on the psychological health of the aged people, the existing literature has not produced consistent and uniform conclusions. Varying interpretations arose due to cryptic complexities in the fertility-health linkage [19]. And large evidence shows a possible bidirectional correlation between elevated fertility and diminished mental health, challenging conventional assumptions [19][20]. Consequently, controlling for endogeneity is imperative to detect this association [18]. Hence, this study adopts an instrumental approach, and addresses two specific questions accordingly.

3 Methods

3.1 Analytical Strategy

Utilizing the 10-item CESD-10, depressive symptoms were assessed during the fourth survey. These items were derived from the original 20-item CESD, eliminating redundancy. Respondents were instructed to indicate their feelings over a last-week period, with values scoring between 0 and 30. The CESD-10, proficiently validated for general groups, demonstrated reliability and validity within our Chinese community-dwelling elderly population [21]. Examining cognitive impairment related to depressive symptoms, scores achieving a minimum of 20 were labeled as depressed patients (D); 10 represented minimal depressive symptoms (DS) while below ten signified presence of none (NDS) [22] .

This current study utilizes biological offspring quantity as a predictor variable while leveraging gender composition of the primary couple's first two offspring to establish predictions. The selected variables are binary, with respondents classified as having identical (daughter-daughter or son-son) or different (daughter-son or son- daughter) sex combinations among their two firstborn offspring. As the second offspring's gender reflects natural randomness, it serves as an outside influence and obviates the necessity for considering collaterally influencing factors.

Due to the complexity created by blending prevalent high fertility rates with significant depressive indicators, conventional multivariate analyses may be insufficient. In lieu of traditional OLS analysis where offspring quantity is hypothesized as external, we have opted

for an instrumental variable methodology [20]. Considering the endogeneity of our key variable (offspring quantity), we require an instrumental indicator. Luckily, we have a unique opportunity in the form of the naturally occurring sex ratio among children which can establish exogeneous IVs.

Following van den Broek and Tosi (2020) [20], We tap into the global trend towards mixed-sex offspring as demonstrated by numerous demographic studies revealing larger odds of having a third birth after two sons or two daughters, compared with one son and one daughter [23][24].

Due to random natural and exogenous factors, couples with same sex offspring might have higher actual fertility rates compared to their counterpart group. Specifically, while certain observed/unobserved prerequisites for mental health issues such as young age at first birth, specific personality traits and socio-economic deprivation may correlate with completed fertility rate, they demonstrate no systematic link with the fertility disparity between two child-birth cohorts distinguished by the child's gender composition. This paves the way for calculating a potential causal influence of elevated fertility frequency. Employing this random assignment, we quantify the causal impact of increased childbearing on mental health through a two-phase methodology. The first-stage model expresses as equation 1:

$$X_i = \alpha_0 + \alpha_1 Z_i + \varepsilon_i \tag{1}$$

In the first stage, an analysis is performed to establish a regression relationship between the children number X for person i and the exogenous instrument Z, i.e. whether or not the two oldest offspring are of identical sex. This model presented in Estimate α_1 denotes the excess number of children that, owing to the preference for gender balance offspring, parents with the two oldest children are of the same sex have over parents with the two oldest children are of different sexes. In this instance, the exogenous distinction in fertility is leveraged in the second stage to evaluate the potential causal impact of procreating additional offspring on mental well-being. In the second stage, the model using IV-Probit regression, denotes that the likelihood of subpar mental status denoted by Y for person i is regressed on $\hat{X}$, which represents the number of children forecasted in the first stage.

$$\Pr(Y_i = 1) = \phi\left(\beta_0 + \beta_1 \hat{X}_i + \mu_i\right) \tag{2}$$

3.2 Analysis of factors influencing depression

In the analysis of influencing factors of depression in the elderly, we focused on a series of potential influencing factors, including education level, living environment, whether there are sons, age, number of living children, and gender. This study aimed to gain insights into the associations between these factors and depressive symptoms in older adults to help develop more effective interventions and supports.

We conducted this analysis using the Spearman correlation coefficient, a nonparametric statistical method that is particularly useful for ratings data or rating data converted from observations of continuous variables. Spearman's correlation coefficient provides a wider range of correlation conditions and is applicable to various data distribution patterns by evaluating the rank-order relationship between variables rather than directly considering their original values.

Specifically, the Spearman correlation coefficient is calculated in a way that can better reflect the nonlinear relationship between factors affecting depression in the elderly. In our study, we looked at multiple dimensions such as educational attainment, living environment, having sons, age, number of living children, and gender. Through the calculation of Spearman correlation coefficients, we were able to evaluate the rank-order relationships between these factors to gain a more comprehensive understanding of their potential links to depression levels in older adults. In the context of Spearman correlation analysis, denoted by ρ, the correlation coefficient is determined by the rank difference (d) and the sample size (n). The Spearman correlation coefficient (ρ) varies between -1 and 1: 0 indicates no correlation, positive values indicate a positive correlation, and negative values indicate a negative correlation. A higher ρ value indicates a stronger correlation, while a lower ρ value indicates a weaker correlation. Specifically, we formalize the Spearman correlation coefficient calculation as:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{3}$$

This analysis method can not only take into account the monotonic relationship between factors, but also handle different types of data, allowing us to more comprehensively understand the complex associations between factors affecting depression in the elderly. The application of this method is expected to provide more targeted directions for future interventions and support measures to improve mental health and quality of life in older adults.

3.3 Transformer-based depression risk prediction method

When developing a Transformer-based prediction model for elderly depression risk, comprehensive data preparation is essential. The input features encompass variables such as educational level, living environment, presence of sons, age, number of living children, and gender. To effectively train and evaluate the model, a high-quality dataset containing pertinent information about elderly individuals and their mental health status is crucial.

In the initial data preparation phase, the focus lies on preprocessing steps, including feature extraction and low-frequency feature filtering. This meticulous preparation ensures that the model can adeptly capture features relevant to predicting depression risk in the elderly while minimizing unnecessary interference.

For feature extraction, we leverage a pre-trained Transformer model to encode and extract rich semantic information. This process encompasses a variety of features such as educational background, living conditions, familial relationships, and demographic details. Our feature extraction involves the use of an encoder, comprising multiple layers, each containing a multi-head self-attention sub-layer and a feed-forward neural network sub-layer.

The self-attention mechanism empowers the model to focus on relevant information from different features, allowing it to allocate diverse semantic information through multiple attention heads. The specific calculations involve intricate self-attention mechanisms, enabling the model to weigh the importance of various input features and effectively learn complex patterns associated with elderly depression risk prediction. The specific calculation is as follows:

$$\text{a}\,(Q, K, V) = \text{softmax}\,(\frac{QK^T}{\sqrt{d_k}})V \tag{4}$$

where Q, K, and V denote the query, key, and value matrices obtained through linear transformation, where d_k signifies the dimensions of each attention head.

Upon the completion of the encoder layer's output passing through the feedforward neural network sub-layer, a merging process occurs through residual connections and layer normalization. This entails combining the outputs of both the multi-head self-attention sub-layer and the feedforward neural network sub-layer:

$$\mathrm{H} = \mathrm{LayerNorm}(\mathrm{X} + \mathrm{FFN}(\mathrm{a})) \tag{5}$$

In the context of the predictive model, denoted as F, the input feature is represented by X. The training phase involves utilizing a training dataset containing actual labels, during which the model parameters undergo fine-tuning. This fine-tuning process aims to minimize the disparity between the predicted outcome F and the authentic labels within the training data. Subsequently, the model undergoes an evaluation phase using an independent set of test data. Various evaluation metrics are then computed to gauge and assess the overall performance of the model.

$$F = \sigma(W \cdot H + b) \tag{6}$$

$$Loss = \frac{1}{2n}\sum_{i=1}^{n}\left(\hat{F}^{(i)} - F^{(i)}\right)^2 \tag{7}$$

Here, $\hat{F}^{(i)}$represents the actual depression risk value, while W and b are learnable parameters. Overall, the Transformer-based depression risk prediction model for the elderly effectively achieves the conversion from the family situation characteristics of the elderly to depression risk prediction by stacking multiple encoder and decoder layers, using self-attention mechanism and feed-forward neural network. To facilitate training and optimization, each sub-layer adopts residual connections and layer normalization. Such a design architecture aims to ensure that the model can efficiently capture important information and improve overall performance.

4 Experiment

4.1 Data

This current study draws on data from China Health and Retirement Longitudinal Study (CHARLS, see http://charls.pku.edu.cn/zh-CN), a nationally representative long-term research project of Chinese adults over 45, featuring community analyses on socioeconomics and health. The fourth national study encompassed a stimulus collection period from May 2011 till March 2012 and involved a total sample size of17,641 participants. This report expanded on the previous concerns by assessing depressive states and cognition. The CHARLS database was utilized via web in October 2020, gaining prompt approval.

For the purpose of our research, we established eligibility criteria including age, demographics, and CESD-10 scale results. From a pool of 17,641 individuals completing the fourth CHARLS follow-up, 3874 met these criteria.

In the current study utilized a cohort of primary respondents over age 60 but with no fewer than two children. As detailed subsequently, the viability of our analytical technique is contingent upon people’s common preferences towards mixed-sex progeny. There were some individuals with missing information about children’ s birth year or the outcome measure, and

also there were some individuals whose second child was not a singleton. All these data is not suitable for the current study, then after eliminating data that not suitable, we obtain 3874 valid samples nested in 17,641 respondents. Since heteroscedasticity arose from nested data, we utilized robust standard errors to assess our models accurately. To ensure our findings are robust, we repeated our analyses whereby one randomly chosen observation dropped from the sample for all 1023 participants in the fourth investigation. These calculations yielded similar outcomes to those presented herein.

4.2 Instrumental Variable: Mixed-gender Preference

The descriptive statistical analysis pertaining to this sampled population was delineated in Table 1. In general, the parent within the sampling population had 3 children. Conforming to prior researches addressing a demonstrated penchant towards birthing mixed-sex offspring such as those reported in some researches [25][26][27], the parents wherein both of their the two oldest offspring were of the same gender exhibited an enhanced total reproductive quantity compared to their counterparts. In addition, this latter cohort exhibited diminished instances of subpar psychological well-being, as measured through an assessment of 10 or above on the CESD-10 scale.

Table 1. Descriptive statistics, CHARLS 2011 (M±SD/%)

Variable	All sample (n =3874)	Two firstborn children of identical sex		Two firstborn children of different sexes	
	M±SD/%	*n*	M±SD/%	*n*	MM±SD/%
Number of children	3.82±1.58	1972	3.91±1.61	1902	3.72±1.53
Suboptimal mental health(CES-D≥20)	42.5%	842	42.7%	803	42.2%
Age	68.85±6.92	1972	68.96±7.02	1902	68.75±6.79

Table 2 illustrates the naïve probit as well as the instrumental variable probit methodologies employed for predicting diminished mental health. The naïve probit analysis identifies a pronounced negative correlation between the quantity of offspring and the risk of subpar mental health. Significantly, as considered previously, the prerequisite of this model implying that quantity of offspring is an external factor might be somewhat exaggerated, and the supposition must take into account that the influence may be overestimated due to unidentified suppressing variables.

The first stage of the instrumental variable model demonstrated that person had the two oldest progeny of similar gender tended to conceive an augmentative number of offspring by approximately 0.19 additional children in comparison to those birthing two oldest progeny with disparate genders. This disparity between those groups was significantly appreciable (F-statistic surpassed 10 notably), thereby indicating primary instrument's effectiveness in predicting total biological descendants through stage one ($F(1, 4898) = 43.69$, $p < .001$). The exogenous fertility disparity across parental groups was employed at stage two to ascertain the

Table 2. Results of two regression models predicting mental health (CES-D ≥20), CHARLS 2011

	Naive probit		IV probit			
	First stage		First stage		Second stage	
	Coef.	(SE)	Coef.	(SE)	Coef.	(SE)
Number of children	-0.08***	0.01			-0.06	0.08
Sex composition two firstborn children						
Identical				0.05		
Different			-0.19***			
Constant	-0.50***	0.05	3.91***	0.02	-0.43	0.24
Wald test of exogeneity: chi2(1) = 0.01				Prob > chi2 = 0.93		

casual impact of offspring quantity upon mental well-being. Per our hypothesis, the model revealed no substantial impact of an extra child on subpar mental health. Evaluation was executed through a Wald test, with statistical relevance defined as $p<0.05$. This demonstrated the predictions of the IV-probit model were closer to reality compared with those of the naïve probit model with non-exogeneity assumptions.

4.3 Depression risk prediction results

As shown in Figure 1, we visualized the distribution of participants' depression levels. A person with a depression level of more than 20 is considered to be suffering from depression. We observe that the abscissa represents the depression score, and the ordinate represents the proportion of the corresponding depression score. One notable feature is that the proportion of participants with depression scores exceeding 20 points is relatively large, which may indicate that some elderly people may suffer from more severe depressive symptoms. We need to further understand the characteristics of this population, such as education level, living environment, family status, etc., to more fully understand the possible influencing factors of depressive problems.

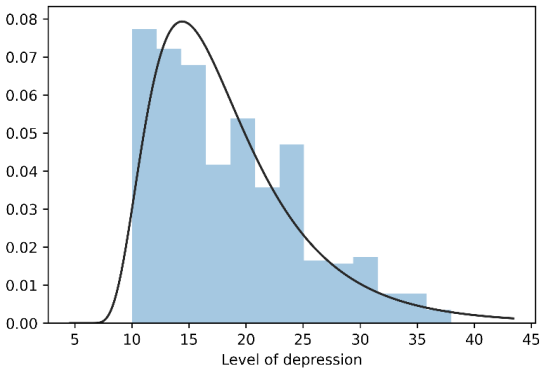

Figure 1 Distribution of depression levels

After conducting Spearman-based analysis of factors influencing depression in the elderly, we obtained a series of Spearman correlation coefficients to measure the degree of association between different factors. To present these results more visually, we can use visualization tools to draw a Spearman correlation coefficient plot (as shown in Figure 2 and Figure 3). For each pair of influencing factors, namely, living environment (K1), gender (K2), education level (K3), age (K4), whether you have a son (K5) and the number of living children (K6), we calculated Pearlman correlation coefficient. The greatest correlation with the degree of depression in the elderly is educational level. The higher the educational level, the less likely it is to be depressed. Additionally, depression levels were least related to age.

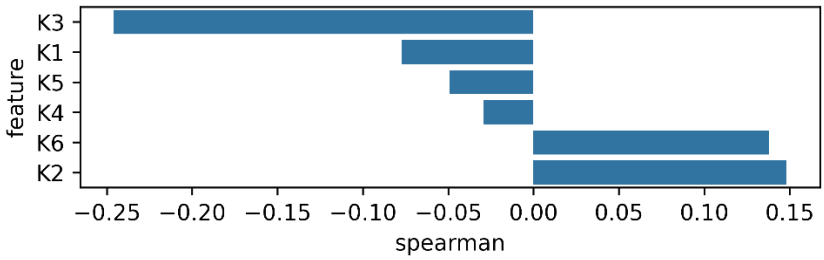

Figure 2 Visualization of Spearman correlation coefficient.

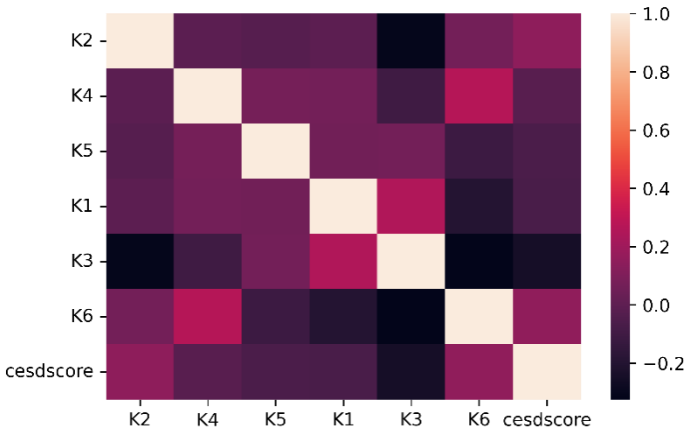

Figure 3 Heat map of depression-related factors.

As shown in Figure 4, by training the Transformer-based depression risk prediction model for the elderly for 300 epochs, we observed that the fitting curves of Loss and RMSE showed a common evolution trend. In the initial stage, as the model learns, Loss and RMSE decrease rapidly, indicating that the model gradually learns the characteristics of the data and improves the fitting ability. After about 180 epochs, these two curves gradually flattened out and formed a stable state. At this point, the model has relatively fully learned the patterns in the data and achieved good performance on the training set and validation set. Further increasing epochs no longer significantly improves performance, showing a convergence trend of the model. This process reflects the learning process of the model, emphasizing that the model has achieved

satisfactory fitting results after 180 epochs, providing a reliable basis for predicting the risk of depression in the elderly.

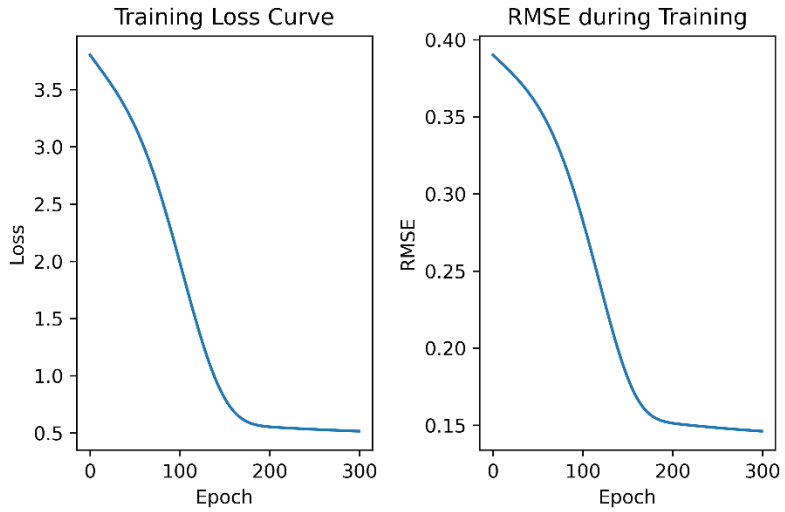

Figure 4 Curve fittings

5 Conclusion

Drawing on CHARLS data of parents aged 60 and older, this study addresses the question of whether the offspring quantity has an impact on the mental well-being of people over the age of 60 in China. Studies investigating the association between offspring quantity and mental well-being in later life often employed predominantly explanatory design, potentially leading to biases related to selection or reverse causation. Utilizing an innovative instrumental variable methodology based on people's preferences for mixed-sex, we were able to ascertain the empirical influence of elevated fertility on mental well-being. The results suggest that additional offspring do not have an effect on mental well-being among Chinese elderly aged above 60 years old. The result suggests that (1) Chinese people have fertility preference for mix-sex; (2) and parents with more offspring had a substantial negative effect on depression measures when endogeneity was not taken into account, but the offspring quantity had a non-substantial effect on parental mental health when estimated using instrumental variables.

This research indicates a tendency in Chinese elderly for mixed-sex offspring, echoing previous researches findings [28][29][30]. Influenced by Confucianism, in terms of the purpose and significance of childbearing, traditional Chinese fertility culture is mainly manifested in the concepts of "raising children to for the sake of old age", "having many children for good fortune" and so on. In agrarian societies, families primarily shoulder societal responsibility for eldercare. So, children always play an important role for elderly care, and daughters also have the irreplaceable utility of sons [31], especially along with the development of society [32], and it is often said that "daughters are the coats of their parents". Therefore, the people of China prefer to have two children, a daughter and a son together to

form a "good" character [33]. So, this result just confirmed previous research which showed that whether in the past or now, having both children is the highest ideal for Chinese people in terms of fertility. This result suggests in a sense that the "two-child" and "three-child" birth policies are in line with Chinese people's birth preferences.

This research validates that the link between offspring quantity and elderly mental well-being yields inconclusive outcomes with instrumental variable utilization. This is consistent with some of the results of Chen (2019) [34] who used data of the Chinese Family Tracking Survey (CFTS) to investigate the correlation between the offspring quantity and depressive symptoms among older adults. Scholars believe that although the old-age resources of parents with a small number of children are objectively more scarce, their better inter-generational relationship with their children and further improvement in social security could offset limitations due to offspring quantity. Song et al. (2011), from the perspective of an only child's ability to age, by comparing the standby interaction behaviors of only children with those of non-only children and their parents, pointed out that only children tend to reside with parents, and no appreciable distinction was noted between the two groups concerning monetary aid or emotional contact with their parents [35]. Kruk and Reinhold's (2014) [19] instrumentally variable analysis found disparities regarding depressive symptoms among multiparous mothers across several European nations. Notably, depression and loneliness are separate entities.

Utilizing an identical methodology as this present research, Thijs van den Broek (2020) discovered increased offspring alleviated potential psychological distress in aging white mothers above 65 years of age [20]. Thijs van den Broek's results contrast with ours possibly due to China's social transformation. Notably, its substantial internal migration over recent decades is corroborated by the Third, Sixth, and Seventh National Censuses, that the floating population was witnessed a significant escalation from 6.57 million in 1982 to 221.43 million in 2010, and subsequently reaching 375.82 million by the year of 2020. There are two kinds of migration-- rural-to-urban labor migration augmented with increasingly educated young adults relocating to major metropolises [36], with the main body of the migration aged between 22 to 44 years old, which accounts for more than 70% of the total [37]. And this directly led to the widening of intergenerational spatial distance which constrained the availability of intergenerational support for the elderly. Shu & Tong (2017) analyzed from the perspective of intergenerational separation, and found that separation of adult children from their parents leads to a decrease in intergenerational support [38]. Certainly it could have a considerable influence on the elderly' mental well-being [39]. Secondly, the development of the economy and the social transition might undermine the filial culture as mentioned above. As a result, the weakening of intergenerational relationships and the alienation of kinship ties surely impacted the traditional family function of old age care.

Although our instrumental variable methodology guarantees bias-free results, it offers limited precision, resulting in wide confidence intervals for coefficients. Consequently, the value assigned to each additional child using this method appears statistically indifferent between zero and its sizable negative counterpart derived through classic OLS regression. Therefore, concluding on potential protective effects from additional children for Chinese parents is challenged by our founding. Moreover, this study utilizes an instrument that reveals old Chinese parents' inclination towards planning another child if they had no girl or boy amongst their two oldest offspring. It's important to note that it doesn't anticipate surprise pregnancies,

which can potentially trigger persistent psychological issues into adulthood [40]. Consequently, the conceivable impact of further children on late-life mental health as explored via these instrumental variables is invalid for unanticipated births.

Despite these drawbacks, it's noteworthy that an additional child does not wholly influence mental well-being for Chinese parents. While, currently in China, to tackle the rapidly aging population and alleviate pressures on elderly single-child households, fertility policies are progressively being relaxed. But, from the results of this study, it is clear that additional offspring alone won't rectify the core problem of family aging, rather, a policy-driven enhancement in child-parent support quality is necessary. The relevant support policies for children to perform filial piety should be improved, such as the implementation of a more humanized and flexible system of family leave and also living arrangements concerning children with their parents, which will exponentially amplify both the temporal and spatial availability for children to attend to eldercare matters, providing the necessary daily attention and emotional comfort for their elderly parents.

In addition, this article introduces a prediction method based on Transformer, which uses a powerful self-attention mechanism and multi-head self-attention to more accurately capture the diverse factors of depression risk in the elderly, providing a new and reliable tool for personalized intervention, and is expected to comprehensively solve the problem of depression in the elderly. For community management, this research hopes to help government to pay attention for the elderly besides children.

References

[1] Yao Y. (2001). A review of the researches on providing for the aged in household in china. *Population & Economics*. (2001) 01-0033-11

[2] Liu Susu, Ouyang Zheng, Wang Haitao. (2016). Overview of social relationship research in the elderly: a perspective based on the escort model. *Population & Development.* Vol. 22 No. 5 2016

[3] Zhong Y. (2023). Research on Anxiety about Elderly Care and Its Mitigation Measures in the Context of Population Aging. *Jiangnan Forum.* Nov. 2023

[4] Van den Broek, T., Tosi, M., & Grundy, E. (2019). Offspring and later-life loneliness in Eastern and Western Europe. *Journal of Family Research.* 31(2), 199–215.

[5] Katherine Keenan & Emily Grundy, 2018. Fertility History and Physical and Mental Health Changes in European Older Adults. *Eur J Population*

[6] Karsten Hank & Michael Wagner (2012). Parenthood, Marital Status, and Well-Being in Later Life: *Evidence from SHARE*.Soc Indic Res.

[7] Regina M. Bures, Tanya Koropeckyj-Cox, and Michael Loree. (2009). Childlessness, Parenthood, and Depressive Symptoms Among Middle-Aged and Older Adults. *Journal of Family Issues*. Volume 30, Issue 5

[8] Grundy, E., van den Broek, T., & Keenan, K. (2017). Number of children, partnership status, and later-life depression in eastern and western europe. *The Journals of Gerontology*: Series B, 74(2), 353–363.

[9] Wang Yumeng, Jianxian Xu, Weihang Yao. (2016). A study on the impact of the number of children on the life satisfaction of middle-aged and elderly people - based on the comparison between urban and rural areas and between genders. *Journal of Agricultural and Forestry Economics and*

Management. 15(06), 735-742.
[10] Liu S.L., Hu S.G., Zhang X.M.. (2020). More children, more happiness? The effect of the number of children on the mental status of rural elderly. *China Rural Economy.* 08, 69-84.
[11] Ritesh Mistry, Gregory D. Stevens, Harvinder Sareen, Roberto De Vogli, and Neal Halfon, (2007). Parenting-Related Stressors and Self-Reported Mental Health of Mothers With Young Children. *American Journal of Public Health*. July 2007, Vol 97, No. 7
[12] Damhnat McCann, Rosalind Bull and Tania Winzenberg. (2015). Sleep Deprivation in Parents Caring for Children With Complex Needs at Home: A Mixed Methods Systematic Review. *Journal of Family Nursing*. 2015, Vol. 21(1) 86 –118.
[13] Vanassche, S., Swicegood, G., & Matthijs, K.(2013). Marriage and children as a key to happiness? Cross-national differences in the effects of marital status and children on well-being.*Journal of Happiness Studies*, 14, 501–524.
[14]Matthias Pollmann-Schult. (2014). Parenthood and Life Satisfaction: Why Don't Children Make People Happy? *Journal of Marriage and Family*. 76 (April 2014): 319–336.
[15]Umberson, Debra (1989). Parenting and Well-being: Theory, Measurement and Stage in the Family Life Course. *Journal of Family Issues.* 10(4): 440-462.
[16] Read, S., Grundy, E., & Wolf, D. A. (2011). Fertility history, health, and health changes in later life: A panel study of British women and men born 1923–49. *Population Studies*, 65(2), 201–215.
[17] Qiang Li, Juanhan Dong & Xin Zhang. (2021). The influence of the number and quality of children on parents' self-evaluation of happiness. *Journal of East China Normal University*. Humanities and Social Sciences No. 4，2021.
[18] Dewei Geng, (2013). More children, more blessings? —— The impact of the number of children on the health of parents. *South China Population*. Vol. 28, No. 3 2013, General No. 117
[19] Kruk, K. E., & Reinhold, S. (2014). The effect of children on depression in old age. *Social Science & Medicine.* 100, 1–11.
[20] van den Broek, T. (2020). Is having more children beneficial for mothers' mental health in later life? Causal evidence from the national health and aging trends study. *Aging & Mental Health.* 1–9.
[21] Chen HJ，Mui AC.(2014). Factorial validity of the center for epidemiologic studies depression scale short form in older population in China．*Int Psychogeriatr*. 2014，26(1) : 49-57
[22] Andresen et al.Center for Epidemiological Studies Depression Scale--Short Form. *APA PsycNet Direct*. https://doi.org/10.1037/t09809-000
[23] Andersson, G., Hank, K., Rønsen, M., & Vikat, A. (2006). Gendering family composition: sex preferences for children and childbearing behavior in the nordic countries. *Demography.* 43(2), 255–267
[24] Hank, K., & Kohler, H.-P. (2000). Gender preferences for children in Europe. *Demographic Research.* 2
[25] Karsten Hank, (2007). Parental gender preferences and reproductive behaviour: a review of the recent literature. *Journal of Marriage and Family*. Volume69, Issue1, February 2007, Pages: 157-173
[26] Hou Jiawei, Huang Silin, Xin Ziqiang. (2014). Changes of Population fertility willingness in China: 1980-2011. *Social Sciences in China.* Issue 4, 2014
[27] Song Jian. (2018). The sex ratio at birth in the environment of relaxed birth policy. *Population and Family Planning*. 2018 NO.5
[28] Yin, Yin. (2012). Raising children for old age and mothers with children: sons or daughters? *Population Studies.* 36(06), 100-109
[29] Yang Xue-yan, Li Shu-zhuo, Shang Zi-juan. (2011). Son Preference or Preference for Having Son and Daughter? Chinese Attitudes towards Reproduction and Policy Implications. *Collection of*

Women's Studies. 1004-2563(2011)06-0027-08
[30] Dou Donghui, Luo Mingming, Liu Xiaoocen. (2019). *Journal of Beijing University of Technology* . (Social Science Edition), 19 (6), 28-38.
[31] Zhang, H. (2012). Sons，Daughters and Intergenerational Support. *Population & Development*. Vol. 18 No. 5 2012
[32] Han Jiang wind. (2019). Analysis of the dilemma and prospect of rural daughters supporting their children in the new era. *Legal system and society*. 2019.4
[33] Song, Y. & Yu, L. 2016 Review of the Research on the Fertility Desire and Its Influencing Factors in China:1980-2015. *Northwest Population*. No.1 2017 .
[34] Chen Xinyi. (2022). Number of Children and Later-Life Depression. Doi: 10.27665/d.cnki.gzcju.2022.000222
[35] Song Jian, Huang Fei. (2011). The generational interaction between China's first generation only child and its parents. *Population Studies*. Issue 3, 2011
[36] Jianmei Zhao & Hai Zhong. (2019). A demographic factor as a determinant of migration: what is the effect of sibship size on migration decisions? *Journal of Demographic Economics.* 85, 321–345
[37] Yanling Liu & Jian Feng, (2014). Characteristics and impact factors of Migration in China: Based on the Analysis of the Sixth Census Data. *Human Geography*
[38] Shu, Binbin and Tong Yuying. 2017. Impact of Adult Child Migration on Health of Rural Elderly Parents. (In Chinese) *Population Research*. 41(2)
[39] Zehan Pan & Weizhen Dong, (2020). Can money substitute adult children's absence? Measuring remittances' compensation effect on the health of rural migrants' left-behind elderly parents . *Journal of Rural Studies*
[40] Herd, P., Higgins, J., Sicinski, K., & Merkurieva, I. (2016). The implications of unintended pregnancies for mental health in later life. *American Journal of Public Health.* 106(3), 421–429